

*République Algérienne Démocratique et Populaire*  
**Ministère de l'Enseignement Supérieur et de La Recherche Scientifique**

**Université Mouloud MAMMERY de Tizi-Ouzou**

**Faculté de Génie électrique et Informatique**

**Département Informatique**



# *Mémoire*

*De fin de cycle*

*En vue de l'obtention du diplôme*

*de master en informatique*

## *Thème*

# *Reconnaissance de visages*

*Proposé et dirigé par :*

**M<sup>r</sup> M. Ouamrane .**  
**M<sup>ed</sup> Redha**

*Réalisé par :*

**M<sup>r</sup> BOUHADDID**

2011/2012

## *∞ Remerciements ∞*

*Je tiens vivement à remercier mon promoteur Mr OUAMRANE pour avoir accepté d'encadrer ce travail et aussi pour tous les conseils prodigués qu'il m'a prodigués pour le réaliser.*

*Mes remerciements s'adressent à tous ceux qui m'ont aidé de près ou de loin dans mon travail et qu'ils trouvent l'expression de ma profonde gratitude*

.

# *Dédicaces*

*Je dédie ce modeste travail à :*

*La mémoire de mon très cher père.*

*Ma mère, qu'elle trouve ici l'expression de mon profond  
Amour,*

*Mes sœurs Lamia, Hafida et Ryma,*

*Mon grand père,*

*Mes tantes et leurs familles,*

*Mes oncles et leurs familles,*

*Mes amis ...*

*B. Redha.*

## *Résumé*

La reconnaissance de visages est une technologie biométrique en vogue, elle est très utilisée dans les applications de contrôle d'accès. Dans la littérature, on trouve plusieurs méthodes globales, locales et hybrides de reconnaissance de visages. Le but de ce projet est d'introduire et d'expliquer les approches classiques de reconnaissance à savoir : PCA, LDA, LVQ, SVM, ..., et surtout de présenter les résultats de l'ensemble des tests des différentes méthodes implémenté, des tests qui ont été réalisés sur différentes bases de données et qui ont permis de conclure que lorsqu'on utilise PCA sur une base qui respecte les conditions suivantes : un éclairage uniforme pendant les deux phases d'apprentissage et de reconnaissance, et la prise de différentes poses pour chaque personne pendant la phase d'apprentissage, avec à chaque fois une légère rotation, ceci donne les performances optimales.

# *Sommaire*

<b>Introduction générale.....</b>	<b>1</b>
-----------------------------------	----------

## **Chapitre I : Biométrie et reconnaissance de visages.**

<b>I.1 Introduction.....</b>	<b>2</b>
<b>I.2 La biométrie.....</b>	<b>2</b>
I.2.1 Définition.....	2
I.2.2 Les technologies biométriques.....	3
<b>I.3 La reconnaissance de visages .....</b>	<b>8</b>
I.3.1 Le processus de reconnaissance de visages.....	9
I.3.2 Les classes des techniques de reconnaissance de visages .....	11
1.3.2.1 Méthodes locales.....	11
1.3.2.2 Méthodes globales .....	12
1.3.2.3 Méthodes hybrides.....	12
1.3.3 Les métriques (indices) de performance .....	13
<b>1.4 Conclusion.....</b>	<b>13</b>

## Chapitre II : L'état de l'art de la reconnaissance de visages

<b>II.1 Introduction</b> .....	14
<b>II.2 Analyse en Composantes Principales</b> .....	14
II.2.1 Présentation .....	14
II.2.2 Le processus de reconnaissance (Karhunen-Loeve) .....	15
<b>II.3 Analyse Discriminante Linéaire</b> .....	16
II.3.1 Présentation .....	16
II.3.2 DF-LDA .....	17
<b>II.4 Analyse en Composantes Indépendantes</b> .....	19
<b>II.5 Le réseau de neurones LVQ</b> .....	23
II.5.1 Introduction .....	23
II.5.2 Le réseau de neurones LVQ .....	24
<b>II.6 Machine à vecteurs de support</b> .....	25
II.6.1 Le principe de SVM .....	25
II.6.2 La séparation optimale .....	25
II.6.3 Le concept d'hyperplan .....	26
II.6.4 La classification linéaire .....	26
II.6.5 La classification non linéaire .....	28
II.6.6 Les classificateurs à marge molle .....	28
II.6.7 La classification multi classe .....	29
<b>II.7 Modèle de Markov caché</b> .....	29
<b>II.8 Template matching</b> .....	31
<b>II.9 L'utilisation des points caractéristiques pour la reconnaissance de visages</b> .....	32
II.9.1 Introduction .....	32
II.9.2 Les méthodes Cross Ratio .....	32
II.9.3 La distance De Hausdorff .....	33
II.9.4 Cross Ratio Vs la distance de Hausdorff .....	34
<b>II.10 Appariement élastique des graphes</b> .....	35
<b>II.11 Modèle morphable 3D</b> .....	37
<b>II.12 Conclusion</b> .....	38

## Chapitre III : Evaluation

<b>III.1 Introduction</b> .....	39
<b>III.2 Les différentes bases de données de visage utilisées</b> .....	39

III.2.1 AT&T (ORL) .....	39
III.2.2 UMIST .....	39
III.2.3 HARVARD .....	40
III.2.4 YALE .....	40
III.2.5 UMIST .....	40
<b>III.3 Les performances de la bonne identification (classification) .....</b>	<b>41</b>
III.3.1 Approches Globales et Locales .....	43
III.3.2 Approches Hybrides.....	47
III.3.3 Tableau récapitulatif .....	49
<b>III.4 La méthode adoptée.....</b>	<b>50</b>
III.4.1 PCA Vs LDA.....	50
III.4.1.1 Les principales différences entre PCA et LDA .....	50
III.4.1.2 L'utilisation d'un seuil pour l'élimination des personnes inconnues .....	51
.....	51
III.4.2 PCA Vs SVM .....	57
<b>III.5 L'amélioration des performances de PCA.....</b>	<b>60</b>
III.5.1 Résoudre le problème d'éclairage.....	60
III.5.2 Résoudre le problème de pose.....	61
III.5.3 Les conditions optimales de pose et d'éclairage .....	62
III.5.4 Vérification.....	63
III.5.5 Tableau récapitulatif .....	78
III.5.6 Détermination du nombre de poses à utiliser dans l'apprentissage .....	79
<b>III.6 Conclusion.....</b>	<b>81</b>
<b>Conclusion générale .....</b>	<b>82</b>
<b>Annexe .....</b>	
<b>Bibliographie .....</b>	

## Introduction générale

La reconnaissance de visages humains est considérée actuellement comme l'une des tâches les plus importantes pour relever le défi en reconnaissance des formes. La facilité et la précision avec lesquelles nous pouvons identifier nos amis et nos ennemis, même dans des conditions défavorables, sont des capacités des plus étonnantes du système visuel humain, le but de la reconnaissance de visages est de concevoir des systèmes informatiques capables d'égaliser les êtres humains dans ce domaine.

S'investir dans le domaine de la reconnaissance de visages est sans doute motivé par la multiplicité et la variété des champs d'applications de celui-ci (télésurveillance et vérification d'identités, chirurgie plastique, langage des signes...). L'intérêt pour ce domaine s'explique aussi par le fait que par rapport aux autres filières de la reconnaissance des formes (reconnaissance de l'écriture, de la voix, etc.), la reconnaissance de visages n'a pas atteint un niveau aussi avancé que ces autres filières. Ceci s'explique avant tout par la complexité de la forme en question (le visage) comparée avec les autres formes.

Dans les prochains chapitres, on présentera tout d'abord la biométrie et la technologie de reconnaissance de visages, puis on introduira et on expliquera les algorithmes classiques de reconnaissance de visages, et enfin on donnera les résultats de l'ensemble des tests réalisés sur différentes bases de données et en utilisant différentes méthodes et sous différentes conditions.

## **I.1 Introduction :**

Dans les applications de contrôle d'accès, la biométrie constitue une solution efficace, simple et surtout pas chère, qui assure de bonnes performances. Et parmi toutes les technologies biométriques qui existent, la reconnaissance de visages est l'une des technologies les plus utilisées et les plus adaptées. En première partie de ce chapitre, on introduit la biométrie et ces diverses technologies, et en deuxième partie, on explique les classes et les principes de la reconnaissance de visages.

## **I.2 La biométrie :**

### **I.2.1 Définition :**

La biométrie est “la reconnaissance automatique d'une personne en utilisant des traits distinctifs”. Une autre définition de la biométrie est “toutes caractéristiques physiques ou traits personnels automatiquement mesurables, robustes et distinctives qui peuvent être utilisées pour identifier un individu ou pour vérifier l'identité prétendue d'un individu” [1].

On trouve deux grands axes dans la biométrie : l'identification et la vérification. Avec l'identification ou la reconnaissance, le système biométrique pose et essaye de répondre à la question, “qui est la personne X ? ”. Dans une application d'identification, le dispositif biométrique requiert une information biométrique et la compare avec chaque information stockée dans la base de données, c'est une comparaison un à plusieurs (1:N). Le but des applications d'identification est d'identifier des criminels et des terroristes en utilisant les données des surveillances [1] [2].

Dans la vérification ou l'authentification, le système biométrique demande à l'utilisateur son identité et essaye de répondre à la question, “est-ce la personne X? ”. Dans une application de vérification l'utilisateur annonce son identité par l'intermédiaire d'un mot de passe, d'un numéro d'identification, d'un nom d'utilisateur, ou toute combinaison des trois.

Le système sollicite également une information biométrique provenant de l'utilisateur, et compare la donnée caractéristique obtenue à partir de l'information entrée, avec la donnée enregistrée correspondante à l'identité prétendue, c'est une comparaison un à un (1:1). Le système trouvera ou ne trouvera pas d'appariement entre les deux.

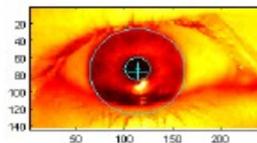
La vérification est communément employée dans des applications de contrôle d'accès et de paiement par authentification [1] [2].

La biométrie offre beaucoup plus d'avantages que les méthodes existantes d'authentification personnelle telles que les clefs, les numéros d'identification (ID), les mots de passe et les cartes magnétiques. En effet elle fournit encore plus de sûreté et de convenance ce qui engendre d'énormes avantages économiques et elle comble les grandes failles de sécurité des mots de passe, surtout avec les facilités actuelles d'accomplir des attaques et de faire du crackage [2].

### **I.2.2 Les technologies biométriques :**

Il existe plusieurs techniques biométriques utilisées dans plusieurs applications et secteurs, et qui exploitent divers informations biométriques à savoir : l'empreinte digitale, le visage, la main, l'iris, la voix, la signature ...

**L'iris :** après l'avoir localisé, on prend des photos en noir et blanc, on utilise ensuite des coordonnées polaires et on cherche les transformées en ondelettes, pour avoir finalement un code représentatif de l'iris. Et on utilise la distance de hamming comme mesure de similarité. La reconnaissance par iris est très utilisée dans les applications d'identification et de vérification, car il est hautement distinctif, unique, sa forme est stable et il est protégé et très robuste, toutefois les équipements d'acquisition coûtent chères [1] [2].



***FIG.I.1-Photo d'iris***

**La rétine :** on mesure dans cette technique la forme des vaisseaux sanguins qui se trouvent dans l'arrière de l'œil. Le dispositif utilisé emploie une source lumineuse rayonnante en direction de l'œil de l'utilisateur, qui doit se tenir fixe devant le dispositif. Et puisque les utilisateurs perçoivent cette technologie comme étant intrusive, la reconnaissance par rétine reste peu populaire et actuellement il n'existe aucun dispositif commercial [1].



*FIG.I.2-Photo de rétine*

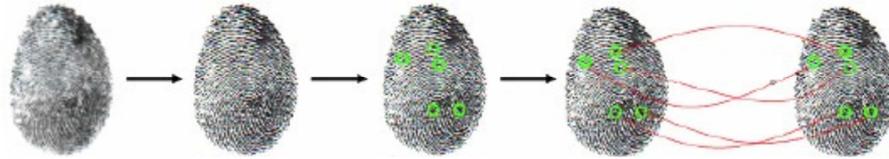
**La voix :** la reconnaissance par voix utilise les caractéristiques vocales pour identifier les personnes en utilisant des phrases mot de passe "pass-phrase". Un téléphone ou un microphone peut être utilisé comme dispositif d'acquisition, ce qui rend cette technologie relativement économique et facilement réalisable, cependant elle peut être perturbée par des facteurs extérieurs comme le bruit de fond [1].



*FIG.I.3-Spectre d'un signal voix*

**L'empreinte digitale :** la reconnaissance par empreinte digitale se base sur le fait que chaque personne a des empreintes uniques. Après la capture de l'image de l'empreinte, on fait un rehaussement de l'image. Ensuite on identifie et on extrait les minuties, qui vont être comparées avec l'ensemble des minuties sauvegardées des autres utilisateurs. C'est l'une des technologies biométriques les plus étudiées et les plus utilisées, surtout dans le contrôle d'accès. Les avantages de cette technique sont : le non changement des empreintes, la haute stabilité et fiabilité, en plus du prix accessible des équipements.

Et les inconvénients sont : la fragilité par rapport aux bruits et aux saletés et la possibilité d'endommager ou d'élimination des empreintes [1] [2].



**FIG. I.4- Le processus de reconnaissance par empreinte digitale**

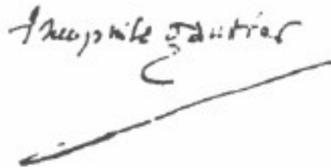
**Les veines :** les techniques basées sur les veines ont été développées en se basant sur le fait que chaque personne a des formes de veines différentes sous la peau [1].

**La géométrie de la main :** la géométrie de la main ou du doigt est une mesure automatisée de plusieurs dimensions, notamment la largeur de la main et des doigts et la longueur des doigts. C'est une technologie qui est rapide et bien développée et qui est facilement acceptée par les utilisateurs. Elle est satisfaisante dans plusieurs situations et elle est appropriée à l'authentification. Néanmoins elle n'est ni trop distinctive ni unique, ce qui la rend inadaptée pour des applications d'identification. Elle offre un taux d'erreur relativement haut et elle n'est pas utilisable avec des personnes jeunes ou âgées [1].



**FIG. I.5- Dispositif de reconnaissance par géométrie de la main**

**La signature** : la vérification par signature est une méthode automatique de mesure des signatures des personnes. Cette technologie examine un ensemble de dynamiques comme la vitesse, la direction, et la pression de l'écriture, le temps pendant lequel le stylo est en contact avec le papier, le temps pris pour faire la signature et les positions où le stylo est relevé et abaissé sur le papier [1].

A handwritten signature in cursive script, reading "Thierry Gauthier". Below the signature is a long, thin, slightly curved horizontal line, possibly representing a stylized underline or a specific stroke from the signature process.

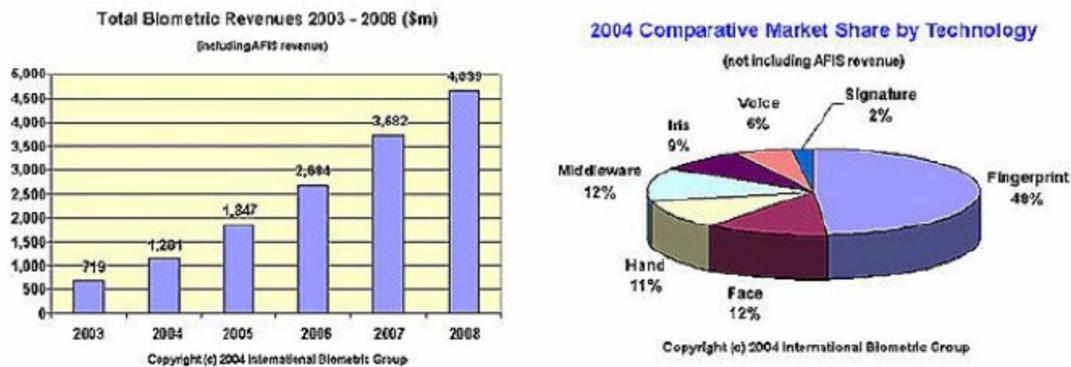
**FIG. 1.6- Signature**

**Le visage** : la reconnaissance basée sur le visage vérifie l'identité, en comparant les données caractéristiques extraites des visages des différentes personnes à partir de vidéos ou d'images fixes. La fabrication des caméras connaît actuellement d'énormes développements, on est capable de prendre des photos de personnes distantes sans qu'ils ne sachent qu'ils sont observés. C'est pour cette raison que la reconnaissance de visages est trop utilisée dans les applications de contrôle de frontières, dans la sécurité des établissements et des zones urbaines et dans l'identification des conducteurs. C'est une technique commune, populaire, simple et qui a beaucoup d'avantages à savoir : l'utilisation des visages qui sont des données publiques, la possibilité de s'intégrer aux systèmes de surveillance existants et elle ne nécessite pas des équipements chers. Cependant, des éléments externes comme l'éclairage, le bruit, l'expression faciale et la posture dégradent les taux de reconnaissance [1] [2].

La reconnaissance de visages est la technologie biométrique qu'on utilisera dans notre système de reconnaissance.

Technologie Biométrique	Précision	Fiabilité	Acceptabilité Sociale	Vitesse (sec)
ID facial	Elevé	Elevé	Très élevé	1.5
Empreinte digital	Très élevé	Modérée	Modérée	6.0
Géométrie de la main	Elevé	Modérée	Modérée	5.0-15.0
Scan de la rétine	Très élevé	Modérée	Faible	5.0-15.0
Reconnaissance vocale	Modérée	Modérée	Très élevé	10
Comparaison des signatures	Modérée	Modérée	Elevé	3.0-5.0

**TAB.I.1-comparaison entre les technologies biométriques**



Technologies	Horizontal Applications	Key Vertical Markets
Fingerprint	Civil Identification	Government Sector
Facial Recognition	Surveillance and Screening	Travel and Transportation
Hand Geometry	PC / Network Security	Financial Sector
Middleware	Retail / ATM / Point of Sale	Health Care
Iris Recognition	eCommerce / Telephony Authentication	Law Enforcement
Voice Verification	Access Control / Attendance	
Signature Verification	Criminal Identification	
Multimedial Biometrics		
AFIS/Live-Scan		

( d'après International Biometric Group)

**FIG.I.7- le visage occupe la deuxième place en biométrie**

### **I.3 La reconnaissance de visages :**

Plusieurs méthodes de reconnaissance de visages ont été proposées durant ces 30 dernières années, suivant deux grands axes : la reconnaissance à partir d'images fixes et la reconnaissance à partir de séquence d'images (vidéo).

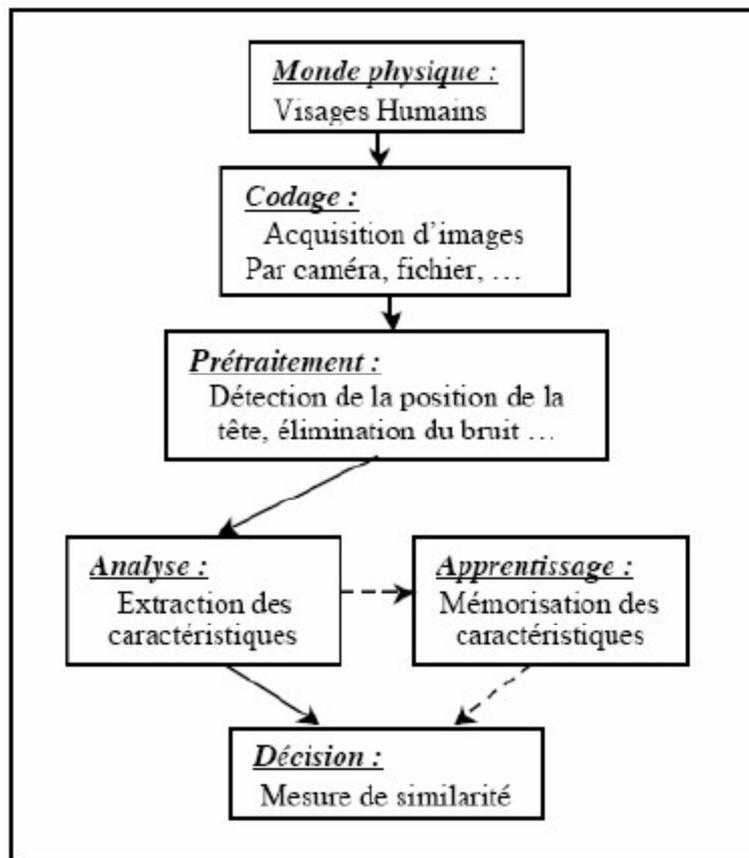
La reconnaissance de visages basée sur la vidéo est préférable à celle basée sur des images fixes, puisque l'utilisation simultanée des informations temporelles et spatiales aide dans la reconnaissance. Dans ce projet, on sait focaliser sur la reconnaissance basée sur les images fixes, puisqu'elle représente la base de tous les travaux.

La reconnaissance de visages est un gros challenge tellement intéressant, qu'il a attiré les chercheurs des différents domaines : psychologie, identification de modèles, réseaux de neurones, vision d'ordinateur, infographie ...

C'est pour cela, que la littérature est si vaste et si diverse. On peut repartir l'ensemble des techniques de reconnaissance de visages basées sur les images fixes, en trois grandes catégories : les méthodes globales, les méthodes locales et les méthodes hybrides.

### I.3.1 Le processus de reconnaissance de visages :

Tout processus automatique de reconnaissance de visages doit prendre en compte plusieurs facteurs qui contribuent à la complexité de sa tâche, car le visage est une entité dynamique qui change constamment sous l'influence de plusieurs facteurs.



**FIG.I.8-Le schéma général d'un système de reconnaissance de visages**

Dans le monde physique, il y a trois paramètres à considérer : L'éclairage, la variation de posture et l'échelle. La variation de l'un de ces trois paramètres peut conduire à une distance entre deux images du même individu, supérieure à celle séparant deux images de deux individus différents [3].

Le codage consiste en l'acquisition d'image et sa digitalisation, il comporte un risque de bruit et donne lieu à une représentation 2-D (la matrice des niveaux de gris) pour un objet 3-D (le visage) [3].

Dans le prétraitement il faut éliminer le bruit par des techniques de traitement et de restauration d'images et procéder à une détection de visages, cette opération est très complexe, surtout dans le cas où l'image contient plusieurs visages ou le cas où l'arrière-plan n'est pas neutre. La restauration d'images ou l'élimination du bruit consiste à compenser les dégradations connues ou estimées et rétablir la qualité initiale de l'image [3].

Les performances globales de tout système automatique de reconnaissance dépendent amplement des performances de la détection de visages [2]. Dans l'étape de détection, on identifie et on localise le visage dans l'image acquise au départ, indépendamment de la position, de l'échelle, de l'orientation et de l'éclairage. C'est un problème de classification où on assigne l'image à la classe visage ou à la classe non visage. On peut diviser les approches de détection en quatre catégories : les méthodes basées sur la connaissance où on code la connaissance humaine du visage, les méthodes de correspondance de masques, les méthodes à caractéristiques invariables où on utilise la couleur, les textures et les contours, et finalement les méthodes les plus répandues et qui sont ceux basées sur l'apprentissage ou les statistiques comme PCA, SVM et Graph matching [4] [5].

Dans l'étape analyse (appelée aussi indexation, représentation, modélisation ou extraction de caractéristiques), il faut extraire de l'image les informations qui seront sauvegardées en mémoire pour être utilisées plus tard dans la phase décision. Le choix de ces informations utiles revient à établir un modèle pour le visage, elles doivent être discriminantes et non redondantes [3].

L'apprentissage consiste à mémoriser les représentations calculées dans la phase analyse pour les individus connus. Généralement les deux étapes d'analyse et d'apprentissage sont confondues et regroupées en une seule étape [3].

La décision : Pour estimer la différence entre deux images, il faut introduire une mesure de similarité [3].

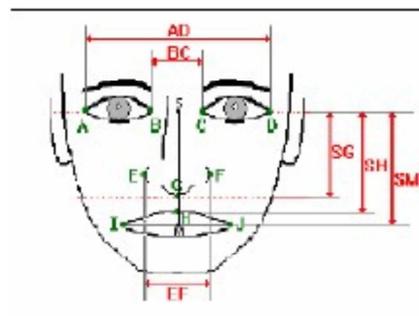
### I.3.2 Les classes des techniques de reconnaissance de visages :

#### I.3.2.1 Méthodes locales :

Ce sont des méthodes Géométriques, on les appelle aussi les méthodes à traits, à caractéristiques locales, ou analytiques.

L'analyse du visage humain est donnée par la description individuelle de ses parties et de leurs relations. Ce modèle correspond à la manière avec laquelle l'être humain perçoit le visage, c'est à dire, à nos notions de traits de visage et de parties comme les yeux, le nez, la bouche, etc. La plus part des travaux réalisés se sont concentrés sur l'extraction des traits à partir d'une image du visage et sur la définition d'un modèle adéquat pour représenter ce visage. Un certain nombre de stratégies automatiques et semi-automatiques ont modélisé et classé les visages sur la base de distances normalisées et angles entre points caractéristiques. La stratégie est qualifiée d'automatique ou non selon que la phase d'extraction des points est faite par l'ordinateur ou qu'elle est assistée par un opérateur. Cette phase constitue l'étape clé du processus, car la performance du système entier dépend de la précision avec laquelle les informations utiles sont extraites [3] [6] [7].

L'avantage de ces méthodes réside dans la prise en compte de la particularité du visage en tant que forme naturelle à reconnaître, en exploitant les résultats de la recherche en neuropsychologie et psychologie cognitive sur le système visuel humain. La difficulté éprouvée lors de la considération de plusieurs vues du visage, ainsi que le manque de précision dans la phase "extraction" des points, constituent leur inconvénient majeur. En plus ces techniques sont facilement affectées par l'information non pertinente.

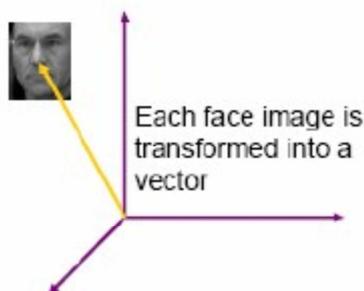


**FIG.I.9-Distances entre points caractéristiques**

Dans cette catégorie, on trouve plusieurs méthodes comme: Pure geometry methods, Dynamic link architecture, HMM...

### **I.3.2.2 Méthodes globales :**

Les méthodes holistiques appelées aussi méthodes globales, sont des méthodes qui sont normalement utilisées dans la reconnaissance de visage, ces méthodes utilisent la région entière du visage comme entrée à l'algorithme de reconnaissance. Ce sont des techniques très réussies et bien étudiées. Ces méthodes offrent les meilleures performances, mais le problème de stockage des informations extraites lors de la phase "apprentissage" reste l'inconvénient majeur [3] [6] [7].



**FIG.I.10-L'utilisation de la région entière du visage comme entrée aux méthodes globales**

Parmi les méthodes les plus connues il y'a: PCA, LDA/FLD, PDBNN...

### **I.3.2.3 Méthodes hybrides :**

Elles combinent les deux types de méthodes, offrant ainsi potentiellement le meilleur des deux types de méthodes.

Elles se basent sur le fait que le système de perception humain emploie les caractéristiques locales et la région entière du visage pour l'identification [6] [7].

Dans cette catégorie on peut citer les méthodes suivantes: Hybrid LFA, Shape-normalized, Component-based...

### **I.3.3 Les métriques (indices) de performance :**

L'exactitude de tout système biométrique est représentée par deux taux d'erreurs qui se produisent. Le système biométrique fait 1 des 4 décisions possibles pendant la phase "décision" :

Le véritable sujet est accepté.

Le véritable sujet est rejeté (fre).

L'imposteur est accepté (fae).

L'imposteur est rejeté.

La performance des systèmes biométriques est mesurée par leur exactitude dans l'identification, qui est calculée en utilisant les fre "false rejection error" et fae "false acceptance error". Le taux où on a la valeur du FAR est égale à la valeur du FRR s'appelle le taux d'erreur égal "EER" et il est utilisé pour comparer les systèmes. Souvent, le taux véritable d'acceptation "gar" est également employé comme seule mesure de performance.

Dans tout système, on ne peut pas simultanément diminuer à la fois FAR et FRR, c'est pour cela que lors de la conception du système, on essaie de dégager un seuil de décision. Et puisque la sécurité est habituellement l'objectif principal, un bas FAR est habituellement choisi aux dépens d'un haut FRR [8] [9].

### **I.4 Conclusion :**

La reconnaissance de visages est une technologie biométrique qui est trop utilisée dans les applications de contrôle de frontières, dans la sécurité des établissements et des zones urbaines et dans l'identification des conducteurs. C'est une technique commune, populaire, simple et qui offre beaucoup d'avantages.

## II.1 Introduction

Dans ce chapitre, on introduit et on explique les algorithmes classiques de reconnaissance de visages, notamment : PCA, LDA, LVQ, SVM, HMM, Template matching, EGM...

## II.2 Analyse en Composantes Principales

### II.2.1 Présentation

L'analyse en composantes principales (PCA) est une méthode mathématique qui peut être utilisée pour simplifier un ensemble de données, en réduisant sa dimension. Elle est utilisée pour représenter efficacement les images de visages, qui peuvent être approximativement reconstruites à partir d'un petit ensemble de poids et d'une image de visage standard (*Eigenpicture*). Ces poids sont obtenus en projetant l'image dans un espace de visage engendré par les visages propres (*Eigenfaces*). Ainsi au lieu de stocker un vecteur image de taille Résolution  $R$ , on stocke un vecteur de taille  $N \leq \text{nombre d'image d'apprentissage}$   $M \ll R$ . Dans l'approche PCA la normalisation d'éclairage est toujours indispensable [10].

Les visages propres sont des images de la même taille que les images d'apprentissage et qui montrent des visages ayant un aspect fantomatique. Mathématiquement, elles sont les composantes principales de la distribution des visages, ou les vecteurs propres (*Eigenvectors*) de la matrice de covariance de l'ensemble des images de visage. Chaque image de visage de l'ensemble d'apprentissage peut être exactement représentée en terme de combinaison linéaire des *Eigenfaces* et du visage moyen [11].

Le nombre possible d'*Eigenfaces* est égal au nombre d'images de visage dans l'ensemble d'apprentissage. Néanmoins les visages peuvent être approximés, en utilisant seulement les meilleurs *Eigenfaces* (ayant les plus larges valeurs propres qui représentent en fait, la plupart de variance dans l'ensemble d'images de visage) qui définissent le "*Low Dimensional Space*", permettant ainsi de diminuer les calculs.

Dans le cas où on ne prend que les  $N$  ( $N < M$ ) grandes valeurs propres (*Eigenvalues*) et les vecteurs propres qui leur sont associés, on réduit encore plus la mémoire de stockage et le temps de calcul, mais on réduit aussi les performances dû à la négligence d'une partie de l'information. C'est pour cela qu'on a pris  $N = M$ , bénéficiant à la fois d'une réduction de l'espace ( $R \rightarrow M$ ) acceptable, et aussi en ayant la totalité de l'information, garantissant ainsi les performances optimales.

Une variante de PCA, plus utilisée dans la communauté du traitement du signal est connue sous le nom de transformée de *Karhunen-Loève* [12].

PCA est une technique rapide, simple et populaire dans l'identification de modèle, c'est l'une des meilleures techniques. Les projections de PCA sont optimales pour la reconstruction d'une base de dimension réduite. Cependant, PCA n'est pas optimisée pour la séparabilité (discrimination) de classe. Une alternative qui est l'analyse discriminante linéaire LDA tient compte de ceci.

### **II.2.2 Le processus de reconnaissance (Karhunen-Loeve)**

Au début de la phase d'apprentissage, on fait l'acquisition, la lecture et la normalisation des images d'apprentissage (de taille  $R$ ). Puis on calcule le visage moyen de ces images. On soustrait ensuite le visage moyen de chaque image d'apprentissage normalisée (ces nouvelles images représentent les colonnes d'une matrice  $S$ ). On fait le calcul de la matrice de covariance qui est définie par la formule  $L=S' * S$ . On calcule ensuite les vecteurs propres  $V$  et les valeurs propres  $D$  de la matrice  $L$ . Puis on passe à l'étape de calcul des visages propres selon la formule  $U = S * V * (\text{abs}(D))^{-0.5}$ . Et finalement on calcule les poids des visages de la base (de taille  $M$ ) en les projetant dans le sous-espace engendré par les visages propres "Face Space".

Pendant la phase de vérification, on fait l'acquisition, la lecture et la normalisation de l'image de vérification (de taille  $R$ ). Puis on soustrait le visage moyen (calculé auparavant lors de la phase d'apprentissage) de l'image de vérification normalisée. Ensuite on calcule le poids de l'image (de taille  $M$ ) en utilisant les visages propres comme une base de projection. Et finalement on utilise la distance euclidienne comme une mesure de similarité [13].

## II.3 Analyse Discriminante Linéaire

### II.3.1 Présentation

L'analyse discriminante linéaire (*LDA*) et les discriminants linéaires de Fisher (*FLD*) relatifs, sont utilisés pour trouver la combinaison linéaire des caractéristiques qui séparent le mieux les classes d'objet ou d'événement. Les combinaisons résultantes peuvent être employées comme classificateur linéaire, ou généralement dans la réduction de caractéristiques avant la classification postérieure.

*LDA* est étroitement lié à *PCA*, du fait que tous les deux recherchent des combinaisons linéaires des variables qui représentent au mieux les données. *LDA* essaye explicitement de modéliser la différence entre les classes des données. *PCA* quant à elle, ne tient pas compte des différences entre les classes.

Chaque visage, qui se compose d'un grand nombre de Pixel, est réduit à un plus petit ensemble de combinaisons linéaires avant la classification.

Chacune des nouvelles dimensions est une combinaison linéaire des valeurs de pixel, qui forment un *template*. Les combinaisons linéaires obtenues en utilisant *FLD* s'appellent les *Fisherfaces*, en analogie avec les *Eigenfaces* [14].

*LDA* est une technique qui cherche les directions qui sont efficaces pour la discrimination entre les données.

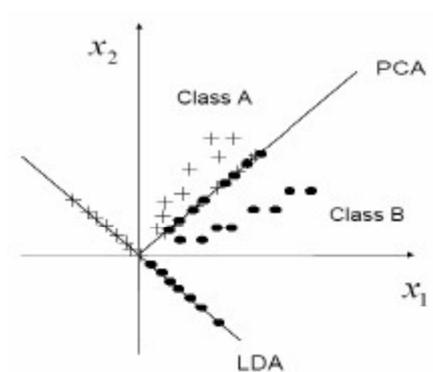


FIG. II.1- Les projections PCA et LDA d'un ensemble de données

LDA est plutôt connue pour sa maximisation de l'éparpillement d'inter-classe "the between-class scatter" et sa réduction au minimum de l'éparpillement d'intra-classe "the within-class scatter", qui se manifeste par le groupement des vecteurs de poids de la même classe (faible distance entre ces vecteurs), et par la séparation des vecteurs de poids de classes différentes (grande distance entre ces vecteurs) [4].

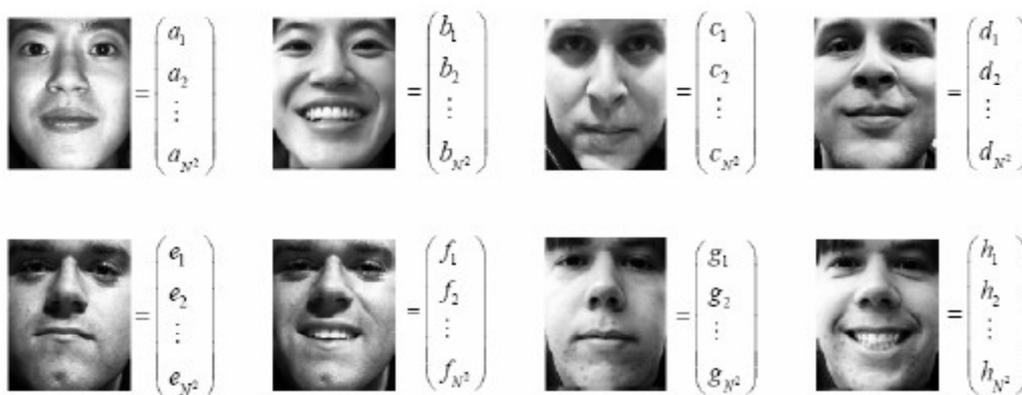
### II.3.2 DF-LDA

La plupart des méthodes traditionnelles basées sur *LDA*, souffrent de l'inconvénient que leurs critères d'optimalité ne sont pas directement liés aux capacités de classification de la représentation obtenue des caractéristiques. D'ailleurs, leur exactitude de classification est affectée par le problème de la petite dimension de l'échantillon "small sample size" qui est souvent rencontré en reconnaissance de visages.

L'algorithme de *DF-LDA* "DIRECT FRACTIONAL-STEP LDA" traite ces deux imperfections d'une façon efficace et rentable, offrant des performances supérieures à celles des techniques de reconnaissance de visages classiques.

Le pseudo code de l'algorithme de *DF-LDA* [15]:

- L'acquisition et la lecture des images d'apprentissage



- La normalisation de ces images
- Le calcul du visage moyen m
- Le calcul du visage moyen de chaque classe

$$\bar{x} = \frac{1}{2} \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_{N^2} + b_{N^2} \end{pmatrix}, \quad \bar{y} = \frac{1}{2} \begin{pmatrix} c_1 + d_1 \\ c_2 + d_2 \\ \vdots \\ c_{N^2} + d_{N^2} \end{pmatrix}$$

$$\bar{z} = \frac{1}{2} \begin{pmatrix} e_1 + f_1 \\ e_2 + f_2 \\ \vdots \\ e_{N^2} + f_{N^2} \end{pmatrix}, \quad \bar{w} = \frac{1}{2} \begin{pmatrix} g_1 + h_1 \\ g_2 + h_2 \\ \vdots \\ g_{N^2} + h_{N^2} \end{pmatrix}$$

- On les soustrait des visages d'apprentissage

$$\bar{a}_m = \begin{pmatrix} a_1 - x_1 \\ a_2 - x_2 \\ \vdots \\ a_{N^2} - x_{N^2} \end{pmatrix}, \quad \bar{b}_m = \begin{pmatrix} b_1 - x_1 \\ b_2 - x_2 \\ \vdots \\ b_{N^2} - x_{N^2} \end{pmatrix}, \quad \bar{c}_m = \begin{pmatrix} c_1 - y_1 \\ c_2 - y_2 \\ \vdots \\ c_{N^2} - y_{N^2} \end{pmatrix}, \quad \bar{d}_m = \begin{pmatrix} d_1 - y_1 \\ d_2 - y_2 \\ \vdots \\ d_{N^2} - y_{N^2} \end{pmatrix}$$

$$\bar{e}_m = \begin{pmatrix} e_1 - z_1 \\ e_2 - z_2 \\ \vdots \\ e_{N^2} - z_{N^2} \end{pmatrix}, \quad \bar{f}_m = \begin{pmatrix} f_1 - z_1 \\ f_2 - z_2 \\ \vdots \\ f_{N^2} - z_{N^2} \end{pmatrix}, \quad \bar{g}_m = \begin{pmatrix} g_1 - w_1 \\ g_2 - w_2 \\ \vdots \\ g_{N^2} - w_{N^2} \end{pmatrix}, \quad \bar{h}_m = \begin{pmatrix} h_1 - w_1 \\ h_2 - w_2 \\ \vdots \\ h_{N^2} - w_{N^2} \end{pmatrix}$$

- La construction des *scatter matrices* (une matrice par classe) :

$$\mathbf{S}_1 = (\bar{a}_m \bar{a}_m^T + \bar{b}_m \bar{b}_m^T), \quad \mathbf{S}_2 = (\bar{c}_m \bar{c}_m^T + \bar{d}_m \bar{d}_m^T),$$

$$\mathbf{S}_3 = (\bar{e}_m \bar{e}_m^T + \bar{f}_m \bar{f}_m^T), \quad \mathbf{S}_4 = (\bar{g}_m \bar{g}_m^T + \bar{h}_m \bar{h}_m^T)$$

- Le calcul de la *within-class scatter matrix* SW:

$$\mathbf{S}_{\text{BTW}} = \mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3 + \mathbf{S}_4$$

- Le calcul de la *weighted between-class scatter matrix*

$$\hat{\mathbf{S}}_{\text{BTW}} = \sum^C \phi_i \phi_i^T, \quad \text{avec } \phi_i = (L_i/L)^{1/2} \sum_{j=1}^C (w(d_{ij}))^{1/2} (\bar{z}_i - \bar{z}_j), \quad \bar{z}_i \text{ est la moyenne de la}$$

classe  $Z_i$ ,  $L_i = \dim(Z_i)$ ,  $L$  est le nombre d'image d'apprentissage,  $C$  est le nombre de classe,  $d_{ij}$  est la distance euclidienne entre les moyennes des classes  $i$  et  $j$ , La fonction pesante  $w(d_{ij})$  est une fonction monotoniquement décroissante de la distance  $d_{ij}$  (il est recommander d'utiliser  $w(d_{ij}) = (d_{ij})^{-2p}$  avec  $p=2,3,\dots$ )

- Le calcul de la matrice  $\mathbf{S}_{TOT} = \mathbf{S}_{WTH} + \hat{\mathbf{S}}_{BTW}$  (la métrique utilisée dans le

$$DF-LDA \text{ est } \Psi = \arg \max_{\Psi} \left( \frac{|\Psi^T \hat{\mathbf{S}}_{BTW} \Psi|}{|\Psi^T \mathbf{S}_{TOT} \Psi|} \right)$$

- L'utilisation des résultats des calculs faits auparavant, dans les 7 étapes principales de l'algorithme :

Step 1. Calculate those eigenvectors of  $\Phi_b^T \Phi_b$  with non-zero eigenvalues:

$$\mathbf{E}_m = [e_1 \dots e_m], \text{ where } m \leq C - 1 \text{ and } \Phi_b \text{ is from } \hat{\mathbf{S}}_{BTW} = \Phi_b \Phi_b^T.$$

Step 2. Calculate the first  $m$  most significant eigenvectors and their corresponding eigenvalues of  $\hat{\mathbf{S}}_{BTW}$  by  $\mathbf{V} = \Phi_b \mathbf{E}_m$  and  $\Lambda_b = \mathbf{V}^T \hat{\mathbf{S}}_{BTW} \mathbf{V}$ .

Step 3. Let  $\mathbf{U} = \mathbf{V} \Lambda_b^{-1/2}$ . Calculate eigenvectors of  $\mathbf{U}^T \mathbf{S}_{TOT} \mathbf{U}$ ,  $\mathbf{P}$ .

Step 4. Optionally discard those eigenvectors in  $\mathbf{P}$  with the largest eigenvalues.

Let  $\mathbf{P}_{M'}$  and  $\Lambda_w$  be the  $M' (\leq m)$  selected eigenvectors and their corresponding eigenvalues.

Step 5. Map all face images  $\{\mathbf{z}_i\}_{i=1}^L$  to the  $M'$ -dimensional subspace spanned by

$$\Gamma = \mathbf{U} \mathbf{P}_{M'} \Lambda_w^{-1/2}, \text{ and have } \{\mathbf{x}_i\}_{i=1}^L, \text{ where } \mathbf{x}_i = \Gamma^T \mathbf{z}_i.$$

Step 6. Further reduce the dimensionality of  $\mathbf{x}_i$  from  $M'$  to  $M$  by performing a

F-LDA on  $\{\mathbf{x}_i\}_{i=1}^L$ , and let  $\mathbf{W}$  (size  $M' \times M$ ) be the bases of the output space.

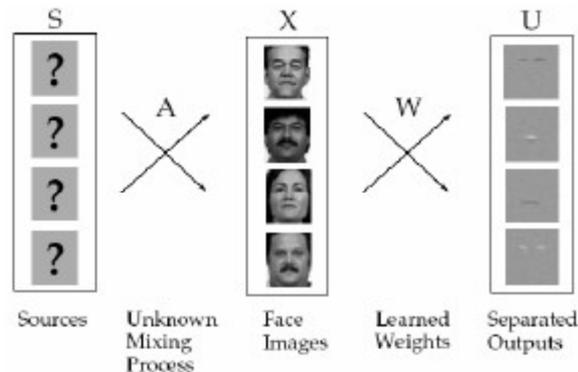
Step 7. The optimal discriminant feature representation of  $\mathbf{z}$  can be obtained

$$\text{by } \mathbf{y} = \varphi(\mathbf{z}) = (\Gamma \mathbf{W})^T \mathbf{z}.$$

## II.4 Analyse en Composantes Indépendantes

*PCA* est une technique optimale de recherche de représentation réduite qui minimise l'erreur de reconstruction, cependant les vecteurs de base tenant compte de l'erreur de reconstruction peuvent ne pas être optimaux pour coder l'information appropriée de l'image à la classification. L'analyse en composantes indépendantes (*ICA*) est une généralisation de *PCA* qui utilise en plus des statistiques d'ordre 2, les statistiques d'ordre plus supérieur, ce qui peut produire une représentation de données plus puissante.

Le but d'ICA est de trouver des vecteurs de base (images de base) localisés dans l'espace et qui sont statistiquement indépendants, en réduisant au minimum la dépendance statistique.



**FIG. II.2- Le modèle de synthèse d'image de l'architecture 1 d'ICA**

Avant l'apprentissage, on soustrait l'image moyenne des images d'apprentissage X et on les filtre par un filtre de blanchiment "whitening filter"  $\mathbf{W}_z = \mathbf{2} * (\mathbf{X}\mathbf{X}^t)^{-1/2}$  pour supprimer les statistiques d'ordre 1 et 2 (la moyenne et les covariances sont mises à zéro et les variances sont égalisées). Pour trouver un ensemble d'images composantes indépendantes, les images d'apprentissage X sont considérées comme étant une combinaison linéaire d'images de base statistiquement indépendantes S, où A est une matrice de mixage inconnue. Les images de base sont retrouvées par une matrice de filtre d'apprentissage  $\mathbf{W}_i = \mathbf{W} * \mathbf{W}_z$ , qui produit des sorties U statistiquement indépendantes. Ces images de base sont considérées comme un ensemble de caractéristiques faciales statistiquement indépendantes, où les valeurs des pixels de chaque image caractéristique sont statistiquement indépendantes des valeurs des pixels des autres images caractéristiques.

$$\text{Face Image} = b_1 * \mathbf{u}_1 + b_2 * \mathbf{u}_2 + \dots + b_n * \mathbf{u}_n$$

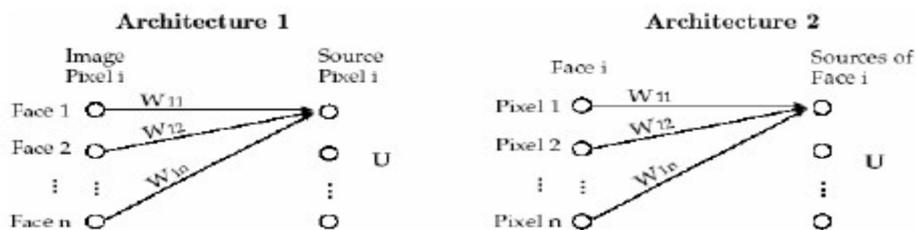
The equation shows a grayscale face image on the left, followed by an equals sign. To the right of the equals sign is a sum of terms:  $b_1 * \mathbf{u}_1 + b_2 * \mathbf{u}_2 + \dots + b_n * \mathbf{u}_n$ . Each  $\mathbf{u}_i$  is represented by a small grayscale image of a facial feature (e.g., eyes, nose, mouth), and each  $b_i$  is a scalar weight.

**FIG. II.3- La représentation d'ICA = (b1, b2, ..., bn)**

La représentation d'ICA est constituée des coefficients  $b$  de la combinaison linéaire des images de base  $U$  qui forment chaque image de visage.

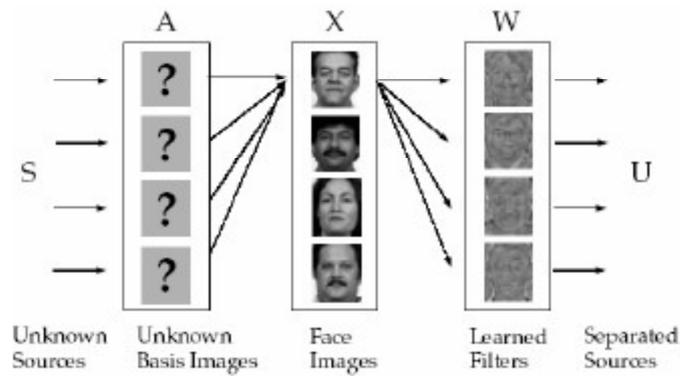
Le nombre de composantes indépendantes extraites par l'algorithme d'ICA est égal à la dimension de l'entrée (le nombre d'images d'apprentissage), et si on veut contrôler ce nombre, au lieu d'exécuter ICA sur les  $n$  images originales, on l'exécute sur un ensemble de  $m$  combinaisons linéaires de ces images, où  $m < n$ . Communément, on choisit pour ces combinaisons linéaires un sous ensemble formé des premières composantes principales de la matrice de donnée, ainsi l'exécution d'ICA sur l'ensemble de ces  $m$  composantes principales produit  $m$  images sources  $U_i$  statistiquement indépendantes.

Dans cette première architecture d'ICA, les images de base  $U_i$  sont statistiquement indépendantes, mais les coefficients  $b_i$  ne le sont pas. Au lieu de séparer les images de visages en ensembles d'images indépendantes, il existe une deuxième architecture d'ICA, dans laquelle les coefficients sont statistiquement indépendants, en séparant les pixels en un ensemble de variables indépendantes codant les images.



**FIG. II.4- Les deux architectures d'ICA :**

*L'architecture 1 pour trouver des images de base statistiquement indépendantes (la séparation des images de visages produit des images composantes indépendantes), et l'architecture 2 pour trouver un code factoriel (la séparation des pixels produit un code factoriel)*



**FIG. II.5- Le modèle de synthèse d'image de l'architecture 2 d'ICA**

Dans l'architecture 2 d'ICA, chaque image de l'ensemble de données est considérée comme étant créée à partir d'un ensemble d'images de base en A, et d'un vecteur de causes d'images "image causes" fondamentales statistiquement indépendantes en S. L'algorithme d'ICA essaye d'inverser les images de base en trouvant un ensemble de filtres  $W_i$  qui produisent des sorties statistiquement indépendantes.

$$\begin{array}{c}
 \text{Face Image} \\
 \hline
 = u_1 * \begin{array}{c} a_1 \\ \text{Basis Image} \end{array} + u_2 * \begin{array}{c} a_2 \\ \text{Basis Image} \end{array} + \dots + u_n * \begin{array}{c} a_n \\ \text{Basis Image} \end{array}
 \end{array}$$

**FIG. II.6- La représentation factorielle d'ICA = (u1, u2, ..., un)**

La représentation factorielle du code est constituée des coefficients indépendants  $u$ , de la combinaison linéaire des images de bases en A qui forment chaque image de visage  $x$  [16].

## II.5 Le réseau de neurones LVQ

### II.5.1 Introduction

L'application des réseaux de neurones artificiels dans la reconnaissance de visages a visé plusieurs problèmes : la classification Homme Femme, la reconnaissance de visages et la classification des expressions faciales [5].

Un réseau de neurones est un système de traitement de l'information qui a été développé comme généralisations des modèles mathématiques assortissant la connaissance humaine. Ils se composent d'un grand nombre d'unités de traitement hautement reliées appelées neurones, travaillant ensemble pour exécuter une tâche de classification donnée.

Un réseau de neurones est un processeur parallèle distribué, ayant une prospérité naturelle pour stocker une connaissance expérimentale. Il ressemble au cerveau humain en trois aspects : la connaissance est acquise par le réseau par un processus d'apprentissage, des forces de connexion reliées ensemble, connues sous le nom de poids synaptiques, sont employées pour stocker la connaissance, et chaque neurone a un état interne appelé seuil ou fonction d'activation (ou fonction de transfert) utilisée pour classifier les vecteurs.

On trouve plusieurs types de réseaux de neurones : l'apprentissage compétitif (non-supervisé, parfois appelé *VQ*), le *LVQ* (apprentissage compétitif supervisé), les cartes auto organisatrices, le perceptron multicouches avec apprentissage par rétro propagation, les réseaux probabilistes, les réseaux à fonctions radiales de base (*FRB*), et les réseaux récurrents [17].

Une classification par réseau de neurones comporte les étapes suivantes :

Tout d'abord une phase de pré traitement des images d'apprentissage et l'association à chaque image d'apprentissage (entrée du réseau) un vecteur de sortie, puis vient l'étape d'initialisation (création des couches du réseau). On fait l'apprentissage (supervisé) du réseau, jusqu'à atteindre une certaine erreur minimale (le réseau apprend à bien classifier les images d'apprentissage). On présente ensuite au réseau une nouvelle image à identifier (phase de reconnaissance ou de simulation ou d'activation du réseau) qui sera finalement affectée à une classe donnée [18].

### II.5.2 Le réseau de neurones LVQ

Parmi tous les types de réseau de neurones qui existent, le type *LVQ* est le type le plus adapté à l'application de reconnaissance de visage, puisqu'au contraire des autres types qui donnent des performances basses, le *LVQ* quand à lui réalise de hautes performances de classification.

L'algorithme de base *LVQ* :

- application d'une entrée
- recherche du neurone vainqueur (sa sortie est égale à 1)
- modification des poids (rapprochement ou éloignement selon que le neurone vainqueur est le neurone désiré ou non)

Dans le langage de programmation *Matlab*, le réseau de type *LVQ* est programmé au moyen de 2 couches : une couche de compétition pour la détermination du neurone vainqueur (la sortie de ce dernier est mise à 1, celle des autres neurones à 0), et une couche de neurones linéaires pour la classification. La 2e couche comporte un neurone par classe, la 1<sup>ère</sup> comporte un neurone par sous-classe.

L'apprentissage modifie les poids de la première couche. Les neurones de la 2<sup>ème</sup> couche combinent certaines sorties des neurones de la 1<sup>ère</sup> au moyen de fonctions OU [17].

## II.6 Machine à vecteurs de support

Une machine à vecteurs de support est une technique de discrimination, c'est une méthode d'apprentissage supervisé utilisée pour la classification et la régression. Elle consiste à séparer deux ou plusieurs ensembles de points par un hyperplan. Selon les cas et la configuration des points, la performance de la machine à vecteurs de support peut être supérieure à celle d'un réseau de neurones ou d'un modèle de mixture gaussienne [19].

### II.6.1 Le principe de SVM

L'idée originale des *SVMs* est basée sur l'utilisation de fonctions noyau "*kernel*", qui permettent une séparation optimale (sans problème d'optimum local) des points du plan en différentes catégories.

La méthode fait appel à un ensemble de données d'apprentissage, qui permet d'établir un hyperplan séparant au mieux les points, dont la dimension est généralement  $n$  [19].

### II.6.2 La séparation optimale

Dans un plan à deux dimensions dans lequel sont répartis deux groupes de points associés à deux classes : les points (+) pour  $y > 0$  et les points (-) pour  $y < 0$ , on peut trouver un séparateur linéaire évident, qui est l'axe des abscisses et qui sépare parfaitement les points. Le problème est dit *linéairement séparable*.

Pour des problèmes plus compliqués, la caractérisation d'un séparateur linéaire peut être très compliquée et tout à fait non optimale. Par exemple dans un plan dans lequel les points (+) sont regroupés en un cercle, avec des points (-) tout autour, aucun séparateur linéaire en deux dimensions ne pourra correctement séparer les groupes. Le problème dans ce cas n'est pas linéairement séparable. La *SVM* utilise les trucs du noyau "*kernel trick*" pour appliquer une classification linéaire à un problème de classification non linéaire [19].

### II.6.3 Le concept d'hyperplan

Afin de remédier au problème de l'absence de séparateur linéaire, le concept des *SVMs* est de reconsidérer le problème dans un espace de dimension supérieure. Dans ce nouvel espace, il existe un séparateur linéaire qui permet de classer au mieux les points dans les deux groupes qui conviennent, avec une distance maximale entre l'hyperplan et les points les plus proches des deux classes, afin de pouvoir classifier plus exactement les nouveaux points, cette distance s'appelle *la marge*. Cet hyperplan est connu sous le nom de *l'hyperplan à marge maximum* ou *l'hyperplan optimal* et les vecteurs (points de dimension  $n$ ) les plus proches de cet hyperplan s'appellent les *vecteurs de support*. On pourra ensuite projeter le séparateur linéaire dans l'espace d'origine pour visualiser le résultat de la classification.

Le séparateur linéaire obtenu peut être un hyperplan, c'est à dire la généralisation à  $n$  dimensions d'une ligne (1D) séparant un espace 2D, ou d'un plan (2D) séparant un espace 3D.

Le changement d'espace se fait au moyen d'une fonction répondant au critère de Mercer. Ce critère permet un changement dans les deux sens, ce qui permet à partir de l'expression de l'hyperplan dans l'espace complexe de classer les éléments dans l'espace de description initial [19].

### II.6.4 La classification linéaire

**On considère un ensemble de données  $\{(x_1, c_1), (x_2, c_2), \dots, (x_p, c_p)\}$  où  $c_p$  est la classe à laquelle appartient le point  $x_p$ ,  $c_p$  est égale à  $(-1)$  ou à  $(+1)$  et  $x_p$  est un vecteur de dimension  $n$  de valeurs appartenant à l'intervalle  $[0,1]$  ou  $[-1,1]$ . Ces données d'apprentissage désignent la classification correcte qu'on voudra que la SVM fasse par la suite, au moyen de l'hyperplan de division, qui obéit à la forme :  $w \cdot x - b = 0$ .**

On s'intéresse aussi aux hyperplans parallèles à l'hyperplan optimal qui sont tangents aux vecteurs de support des classes, et qui sont décrits par les équations :  $w \cdot x - b = 1$  et  $w \cdot x - b = -1$ .

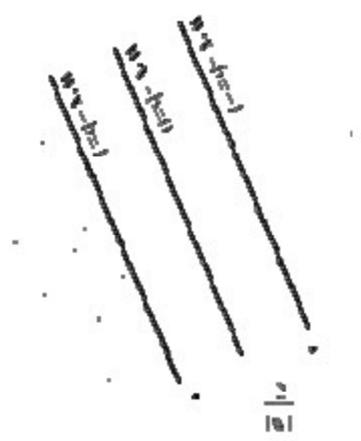
On veut que ces hyperplans maximisent la distance avec l'hyperplan de division et qu'il n'y ait aucun point de données entre les deux. En employant la géométrie, on trouve que la distance entre les hyperplans est  $2 / |w|$ , ainsi le but est de réduire au minimum  $|w|$ .

Pour exclure les points de données, on doit s'assurer que pour chaque  $i$  on a  $w \cdot x_i - b \geq 1$  ou  $w \cdot x_i - b \leq -1$ , ou sous une autre équation  $c_i (w \cdot x_i - b) \geq 1$  avec  $1 \leq i \leq p$ . Le problème maintenant est de minimiser  $|w|$  sous cette dernière contrainte, ceci est un problème d'optimisation de *programme quadratique (QP)*. Il existe plusieurs algorithmes spécialisés pour résoudre rapidement les problèmes *QP* et la méthode la plus commune est l'algorithme *SMO*.

Après l'apprentissage, la *SVM* peut classifier les nouvelles données (données de test) en utilisant la règle de décision suivante :

$$\hat{c} = \begin{cases} 1, & \text{if } w \cdot x + b \geq 0 \\ -1, & \text{if } w \cdot x + b \leq 0 \end{cases}$$

On peut constater d'après la règle de classification qu'elle est seulement en fonction des vecteurs de supports (les données d'apprentissage se trouvant sur la marge) [19].



**FIG. II.7- La classification en deux classes par la SVM**

## II.6.5 La classification non linéaire

L'algorithme original de l'hyperplan optimal est un classificateur linéaire, et pour avoir un classificateur non linéaire on applique le *kernel trick* à l'hyperplan à marge maximum, l'algorithme résultant est formellement semblable, sauf que chaque produit scalaire est remplacé par une fonction noyau non linéaire.

Parmi les noyaux les plus communs, on trouve : le noyau polynomial (homogène)

$K(x, x') = (x \cdot x')^d$ , le noyau polynomial (inhomogène)  $K(x, x') = (x \cdot x' + 1)^d$  la

fonction radiale de base RBF  $K(x, x') = \exp(-\gamma \|x - x'\|^2)$  avec  $\gamma > 0$ , la fonction radiale de

base de type gaussien  $K(x, x') = \exp(-\frac{\|x - x'\|^2}{2\theta^2})$  et la fonction sigmoïde  $K(x,$

$x') = \tanh(kx \cdot x' + c)$  [19].

## II.6.6 Les classificateurs à marge molle

En pratique l'algorithme est instable quand la solution optimale ne sépare pas parfaitement les classes. On a donc fait une modification un peu heuristique, qui consiste à accepter certains exemples comme incorrects, en rajoutant une pénalité dans la fonction de coût, pour la distance entre la position  $x_i$  de l'entrée  $i$  et la marge au-delà de laquelle il devrait se trouver.

Les *classificateurs* à marge molle "*soft margin classifiers*" ont donc comme défaut le choix d'un paramètre  $C$ , qui contrôle en fait la pénalité sur les exemples mal classifiés. On le choisit généralement par validation croisée. Si on prenait une valeur de  $C$  plus grande, ces exemples pourraient devenir des exemples correctement classifiés ou des exemples sur la marge.

Cette méthode a pour équation :  $c_i(w \cdot x_i - b) \geq 1 - \varepsilon_i$   $1 \leq i \leq n$  et le problème d'optimisation devient :  $\min \|w\|^2 + c \sum_i \varepsilon_i$  avec  $c_i(w \cdot x_i - b) \geq 1 - \varepsilon_i$   $1 \leq i \leq n$

Cette contrainte peut être résolue en utilisant des multiplicateurs de Lagrange [19] [20].

## II.6.7 La classification multi classe

Parmi les stratégies adoptées pour résoudre des problèmes de classification en  $q$  classes par SVM, on trouve l'approche *une contre toutes* “*one vs all*” et l'approche par paires “*pairwise*”. On obtient un système de classification multi classe en combinant différents classificateurs SVM (classificateur binaire ou classificateur en deux classes).

Dans la première approche, on fait l'apprentissage de  $q$  SVMs dont chacune sépare une classe de l'ensemble des autres classes.

Dans la deuxième approche, on fait l'apprentissage de  $q(q-1)/2$  SVMs dont chacune sépare une paire de classes. Ces classificateurs sont arrangés en arbre dont les nœuds sont les SVMs.

Une récente comparaison entre plusieurs techniques multi classes favorise l'approche une contre toutes vu sa simplicité et ses bonnes performances de classification. Concernant l'apprentissage, il est préférable d'avoir  $q$  SVMs que d'avoir  $q(q-1)/2$  SVMs et le temps d'exécution des deux stratégies est similaire [21].

## II.7 Modèle de Markov caché

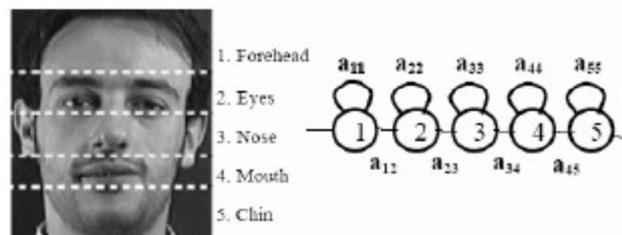
Les modèles de Markov cachés HMMs sont massivement utilisés en reconnaissance de formes, en intelligence artificielle ou encore en traitement automatique du langage naturel.

Les HMMs sont généralement employés pour la modélisation statistique de séries de vecteur de temps non stationnaire. En considérant l'information du visage comme étant une séquence variable dans le temps, le HMM peut être appliqué à la reconnaissance de visages [5].

Les caractéristiques faciales les plus significatives d'une image de visage frontale, à savoir les cheveux, le front, les yeux, le nez et la bouche, se présentent dans un ordre naturel de haut en bas, même dans le cas où il y'a des petites rotations.

En se basant sur cette observation, l'image d'un visage peut être modélisée en utilisant un *HMM* unidimensionnel en assignant à chacune de ces régions un état. Le *HMM* classifie une caractéristique par la propriété de la chaîne de Markov.

Une séquence des valeurs des pixels forme une chaîne de Markov, si la probabilité que le système à l'instant  $n+1$  soit à l'état  $x_{n+1}$  dépend uniquement de la probabilité que le système à l'instant  $n$  soit à l'état  $x_n$ . Dans une chaîne de Markov, la transition d'un état à un autre est probabiliste ce qui donne une distribution de probabilité de l'ensemble des sorties de chaque état. Ce résultat est employé pour faire la comparaison entre deux visages [5] [18].



**FIG. II.8- Les 5 états du HMM (de haut en bas)**

Pour une image de visage d'une personne donnée de l'ensemble d'apprentissage, le but de l'étape d'apprentissage est d'optimiser les paramètres pour mieux décrire l'observation. La reconnaissance est effectuée en assortissant l'image de test contre chaque modèle d'apprentissage (chaque *HMM* représente une personne différente). Et dans la fin de cette procédure, l'image est convertie en une séquence d'observation et la probabilité est calculée pour chaque modèle enregistré. Le modèle avec la probabilité la plus élevée indique l'identité de la personne inconnue [5] [11].

L'approche *HMM* a des performances de reconnaissance satisfaisantes, cependant il nécessite des calculs intenses ce qui augmente les temps d'exécution [5].

## II.8 Template matching

Le template matching est une technique globale de reconnaissance de visage. La corrélation est généralement exploitée pour mesurer la similitude entre un template (masque) stocké et l'image à reconnaître. Les templates devraient être délibérément conçus pour couvrir la variété de variations possibles d'image.

Pendant la recherche dans l'image, l'échelle et la rotation devraient également être soigneusement considérées pour accélérer le processus.

Cette technique a aussi été exploitée pour localiser les caractéristiques saillantes "salient features" comme les yeux le nez et la bouche, dans une image de visage. D'affines transformations ont été introduites dans le processus de reconnaissance, lors de changement de pose [22].

L'algorithme du template matching :

On commence par l'acquisition, la lecture et la normalisation des images d'apprentissage. On calcule ensuite le visage moyen de chaque classe de personnes (les templates). Dans l'étape de reconnaissance, on fait l'acquisition, la lecture et la normalisation de l'image de vérification. Puis on calcule les différences entre l'image de vérification et les templates (les différences sont des images). On calcule après les sommes de ces différences (somme des pixels de l'image différence). Et finalement la somme minimale va référencer la classe auquel appartient l'image de vérification.

## II.9 L'utilisation des points caractéristiques pour la reconnaissance de visages

### II.9.1 Introduction

La reconnaissance de visages basée sur les points caractéristiques est un processus de reconnaissance qui utilise les points caractéristiques : les yeux et la bouche comme une entrée aux algorithmes de reconnaissance. Ces algorithmes peuvent se servir de l'information appropriée telle que l'endroit ou les statistiques locales.

Au départ, dans le processus de reconnaissance, on utilisait les rapports de distances entre les points caractéristiques ou ce qu'on appelle les CROSS RATIO, après on a utilisé la distance de Hausdorff comme mesure de similarité [23].

La reconnaissance basée sur les points caractéristiques est très affectée par la bonne extraction de ces points. Tant que la méthode d'extraction est plus fine et plus exacte, tant que le taux de réussite sera meilleur.

### II.9.2 Les méthodes Cross Ratio

Les premières méthodes locales de reconnaissance se basaient sur la largeur de la tête, et sur les distances entre les yeux et la bouche, en calculant des rapports entre ces distances (CROSS RATIO) [7]. Ces méthodes sont facilement affectées par l'information non pertinente.

L'algorithme de reconnaissance en utilisant les Cross Ratio :

L'extraction des points caractéristiques du visage (affectée par l'orientation de la tête).

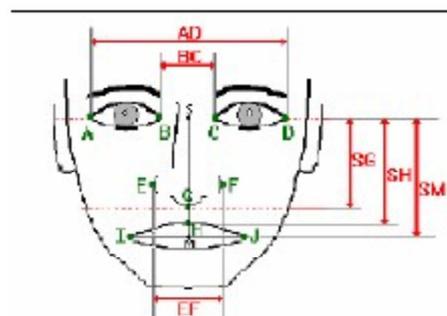


FIG. II.9- Distances entre points caractéristiques

Puis La définition de Cross Ratio pour chaque 4 points d'une ligne (distances invariables). Ensuite on fait la correction de la localisation des points caractéristiques par l'application de la symétrie et du cross ratio. On normalise après le vecteur caractéristique :  $N = F / \|F\|$ , et on utilise finalement la distance euclidienne comme une mesure de similarité.

### II.9.3 La distance De Hausdorff

La distance de Hausdorff est une mesure de distance entre deux polygones ou entre deux figures. La distance de Hausdorff (DH) est régulièrement utilisée en analyse d'image. D'après Rucklidge, elle est considérée comme une mesure de similarité naturelle entre les formes. La distance de Hausdorff du polygone A au polygone B est définie comme :

$$h(A, B) = \max \{ \min \{ d(a, b) \} \}$$
$$a \in A \quad b \in B$$

$d(a, b)$  étant la distance Euclidienne entre les deux points a et b.

$h(A, B)$  est appelée la distance de Hausdorff relative (ou semi distance de Hausdorff).

On peut définir ainsi, la distance de Hausdorff de A et B comme le plus grand des deux nombres  $h(A, B)$  et  $h(B, A)$  :  $DH_{(A, B)} = \max(h(A, B), h(B, A))$  [23] [24].

L'algorithme de reconnaissance en utilisant la distance de Hausdorff :

On prend un ensemble de points caractéristiques, tel que les yeux ou la bouche, de l'image à identifier. On calcule ensuite la distance entre cet ensemble et la totalité des ensembles déjà extraits lors de la phase d'apprentissage. On peut alors continuer à exécuter des transformations sur l'ensemble de points, comme la rotation et la translation, pour affiner les calculs. Si la distance de Hausdorff devient nulle, alors les ensembles de points sont identiques, ou alternativement si la distance de Hausdorff tombe au-dessous d'une certaine valeur seuil, on peut dire que les deux ensembles de points sont semblables.

Puisque la distance de Hausdorff est juste une mesure de distance, on peut imaginer que d'autres mesures de distance, telles qu'Euclidienne ou Procruste, pourrait être employé dans une manière semblable.

## II.9.4 Cross Ratio Vs la distance de Hausdorff

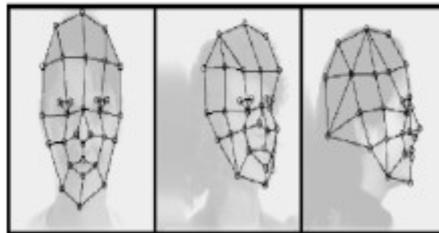
Ce tableau comparatif résume les principales différences entre les deux approches :

	Cross Ratio	la distance de Hausdorff
La rapidité	Plus rapide	Moins rapide
L'échelle (scale)	On peut appliquer l'algorithme à des images de différentes échelles	On utilise des images de même échelle (on peut utiliser des images de différentes échelles, en ajoutant un bloc de traitement d'échelle à l'algorithme)
La performance	Moins performant	Plus performant
L'extraction des points	On doit extraire le même nombre de points de chaque image, ce qu'un algorithme d'extraction ne peut pas assurer surtout avec des images de profile ou avec des images comportant des composantes structurelles (barbe, moustache ...)	On peut calculer la distance entre deux ensembles, de nombre de points différents

**TAB. II.1-** Les principales différences entre les Cross Ratio et la distance de Hausdorff

## II.10 Appariement élastique des graphes

L'algorithme d'EGM "*Elastic Graph Matching*" est décrit en *Architecture Dynamique De Lien DLA*, il représente des visages individuels par un graphe rectangulaire. Chaque nœud du graphe est étiqueté avec un ensemble de *coefficients complexes des ondelettes de Gabor* appelés *jets* différant en orientation et en échelle. Un *jet* est utilisé pour représenter les caractéristiques locales des images de visages en se basant sur la transformée en ondelettes de Gabor. Seulement les magnitudes des coefficients sont utilisées pour la reconnaissance.



**FIG. II.10-** *La représentation des visages par graphe rectangulaire*

Pour la reconnaissance d'un nouveau visage, chaque graphe de la base de données est ajusté à ce nouveau graphe construit, et le bon ajustement indique la personne reconnue.

D'encouragent résultats sont obtenus lors d'utilisation de visages avec de large rotation d'angles. En général la *DLA* est bonne en terme de variation de rotation, cependant le processus de reconnaissance ou d'ajustement est coûteux en calcul.

*L'EGM* utilise la phase *des coefficients complexes des ondelettes de Gabor* pour réaliser une localisation plus précise des nœuds et pour différencier les modèles similaires en leurs coefficients de magnitude.

L'EGM emploie des graphes adaptatifs d'objet "object adaptive graphs", dont les nœuds réfèrent aux points de repère faciaux spécifiques, ou les points fiduciels "fiducial points" dans le visage, à savoir les pupilles, les coins de la bouche et le bout du nez.



**FIG. II.11- Les points fiduciels du visage**

Une nouvelle structure de donnée appelée le graphe de groupe "bunch graph" est introduite pour servir comme une représentation générale des visages, en combinant les jets d'un petit ensemble de visages individuels. Ceci permet au système de trouver les points fiduciels dans un seul processus d'ajustement, réduisant ainsi significativement les calculs.

Pour une image de test, l'algorithme d'EGM essaye de trouver les points fiduciels et d'extraire à partir de l'image un graphe qui maximise la similarité.

L'utilisation des SVMs améliore la performance d'EGM pour l'authentification des visages frontaux.

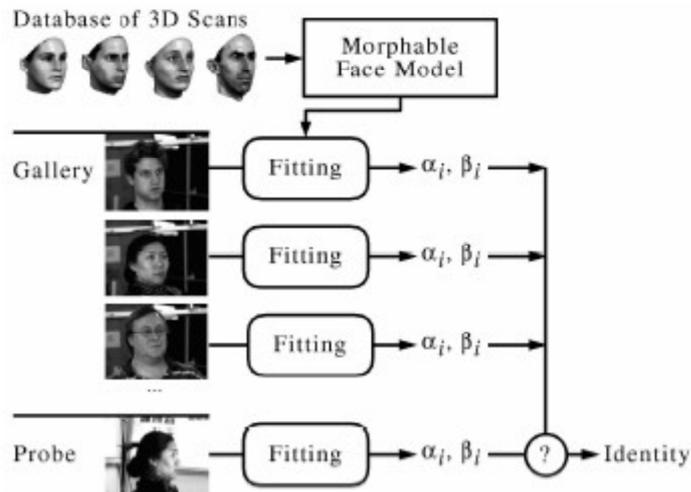
Les avantages d'EBGM "Elastic Bunch Graph Matching" sont : la facilité d'ajouter des données (nouveau sujet, nouveau graphe) en modifiant les graphes déjà existants dans la galerie (échelle, rotation), et la robustesse, puisque l'EBGM est peu sensible aux variations de l'éclairage, de la pose et de l'expression faciale. Néanmoins, l'EBGM nécessite beaucoup de calcul et requiert une grande taille d'image par exemple 128\*128, qui restreint sévèrement l'application dans la vidéo surveillance, où la taille d'image de la zone du visage est très petite [5] [11].

## II.11 Modèle morphable 3D

Le modèle morphable du visage est basé sur une représentation de l'espace de vecteur des visages, le modèle est construit comme toute combinaison convexe des vecteurs de forme et de texture d'un ensemble d'images de visages.

Le raccord du modèle morphable 3D aux images peut être utilisé dans deux approches de reconnaissance de visages. Dans la première approche, après le raccord du modèle, la reconnaissance peut être basée sur les coefficients du modèle, qui représentent la forme et la texture intrinsèques des visages, et qui sont indépendants des conditions de prise d'images. Et dans la deuxième approche, la reconstruction tridimensionnelle du visage peut également être utilisée pour générer des vues synthétiques à partir des images de test et qui seront ensuite transférées à un second système de reconnaissance dépendant du point de vue.

Plus récemment, une technique combine les modèles morphables 3D avec une simulation graphique par ordinateur, de la projection et de l'éclairage. Pour une image d'une personne donnée, l'algorithme estime automatiquement la forme 3D, la texture, et l'ensemble des paramètres appropriés de la scène 3D. Dans ce système, les rotations ou les changements d'éclairage sont des opérations très simples, et tous les éclairages et poses sont couverts par un modèle simple. L'éclairage n'est pas limitée à la réflexion de Lambertian, mais tient compte des réflexions spéculaires et des ombres, et qui ont une influence considérable sur l'aspect de la peau humaine. Cette approche est basée sur un modèle morphable des visages 3D qui capture les propriétés spécifiques à la classe des visages. Ces propriétés sont apprises automatiquement à partir d'un ensemble d'apprentissage formé d'entrées 3D. Le modèle morphable représente les formes et les textures des visages comme des vecteurs dans un espace de visage haut dimensionnel, et il implique une fonction de densité de probabilité des visages naturels dans l'espace de visage. L'algorithme estime automatiquement tous les paramètres de la scène 3D, y compris la position et l'orientation de la tête, la longueur focale de la caméra et la direction de l'éclairage. Ceci est réalisé par une nouvelle procédure d'initialisation qui augmente considérablement la robustesse et la fiabilité du système. La nouvelle initialisation utilise des coordonnées d'image de six à huit points caractéristiques.



**FIG. II.12- Le modèle morphable 3D**

Dérivé d'une base de données de balayages par laser, le modèle morphable 3D du visage est employé pour coder les images d'apprentissage et l'image de test. Pour l'identification, les coefficients du modèle  $\alpha_i$  et  $\beta_i$  de l'image de test sont comparés aux coefficients stockés des images d'apprentissage [11].

## II.12 Conclusion

L'état de l'art de la reconnaissance de visages est très riche, on trouve une multitude d'algorithmes globales, locales et hybrides qui sont utilisés et développés dans les systèmes commerciaux et dans les projets de recherche.

## **III.1 Introduction**

Après avoir survolé la littérature de la reconnaissance de visages, on est passé à une étape d'évaluation de plusieurs méthodes, afin d'aboutir à la méthode la plus optimale et la plus adéquate. Dans cette partie, on présente les résultats de l'ensemble des tests effectués sur différentes bases de données et en utilisant différentes méthodes et sous différentes conditions.

## **III.2 Les différentes bases de données de visage utilisées**

### **III.2.1 AT&T (ORL)**

Cette base contient des images de visage de 40 personnes, avec 10 images pour chacune. Pour la plupart des sujets, les 10 images ont été tirées à différentes heures et avec différentes conditions d'éclairage, mais toujours sur un fond foncé. La taille d'image est 92 x 112.

Les limitations de cette base de données sont : le nombre limité de personnes, la non-conformité des conditions d'éclairage d'une image à une image, et la non assignation de toutes les images par les différentes expressions faciales, rotation de la tête, ou conditions d'éclairage [11].

### **III.2.2 UMIST**

Elle se compose de 1012 images en niveau de gris prises à un éclairage uniforme, de 20 personnes des deux sexes et de diverses races. La taille d'image est environ 220 x 220. De divers angles de pose de chaque personne sont fournis, s'étendant du profil aux vues frontales.

Les limitations de cette base de données sont : aucun angle absolu de pose n'est donné pour chaque image, et aucune information n'est fournie au sujet de l'éclairage utilisé (ni sa direction ni sa température de couleur) [11] [25].

### **III.2.3 HARVARD**

Cette base contient des images de visage en niveau de gris de 10 personnes, avec une soixantaine d'images pour chacune. Les images sont de face et elles ont été tirées à des conditions d'éclairage extrêmement variables. La base est subdivisée en 5 ensembles (Set), où dans chacun, la variation de l'éclairage d'une image à une autre, est différente. La taille d'image est 84 x 96 [26].

Les limitations de cette base de données sont : le nombre limité de personnes, la non disponibilité d'images avec différentes expressions faciales, ou conditions de pose, en plus du non traitement de toutes les conditions d'éclairage.

### **III.2.4 YALE**

Elle se compose de 165 images de visages frontales en niveau de gris de 15 personnes, avec 11 images pour chacune. On trouve 3 angles d'éclairage différents : gauche, centre, et droit, et il existe des images avec lunettes et sans lunettes. La base offre des images incluant différentes expressions faciales : normal, heureux, triste, somnolent, étonnant, et clignotement de l'œil.

Les limitations de cette base de données sont : le nombre limité de personnes, les positions exactes des sources d'éclairage ne sont pas indiquées, il n'y a aucune variation d'angle de pose et les facteurs environnementaux (tels que la présence ou l'absence de la lumière ambiante) ne sont pas également décrits [11].

### **III.2.5 FERET**

Cette base de données contient 11338 images de visages en couleur. Elles ont été rassemblées en photographiant 994 sujets à de divers angles, durant 15 sessions entre 1993 et 1996. Cette base de données est en grande partie une version couleur de la base de données originale de FERET qui contenait 14051 images en niveau de gris. Elle a été créée pour développer, tester et évaluer des algorithmes de reconnaissance de visages.

Les images couleurs ont une résolution : 512 x 768, en format PPM.

13 différentes poses sont disponibles dans la base :

fa (image frontale régulière),

fb (image frontale alternative, prise peu de temps après l'image fa correspondante),

pl (profile de gauche),

hl (demi gauche - la tête est tournée d'environ 67.5 degrés à gauche),

ql (quart de gauche - la tête est tournée d'environ 22.5 degrés à gauche),

pr (profile de droite),

hr (demi droite - la tête est tournée d'environ 67.5 degrés à droite),

qr (quart de droite - la tête est tournée d'environ 22.5 degrés à droite),

ra (image aléatoire - la tête est tournée d'environ 45 degrés à gauche),

rb (image aléatoire - la tête est tournée d'environ 15 degrés à gauche),

rc (image aléatoire - la tête est tournée d'environ 15 degrés à droite),

rd (image aléatoire - la tête est tournée d'environ 45 degrés à droite),

re (image aléatoire - la tête est tournée d'environ 75 degrés à droite) [27] [28].

Les limitations de cette base : elle ne fournit pas une très grande variété de variations de pose, il n'y a aucune information sur l'éclairage employé en plus du très peu nombre d'images par pose pour chaque personne [11].

### **III.3 Les performances de la bonne identification (classification)**

(Dans ce qui va suivre ci-dessous tous les tests ont été fait par l'équipe de recherche de l'IRIToulouse [37])

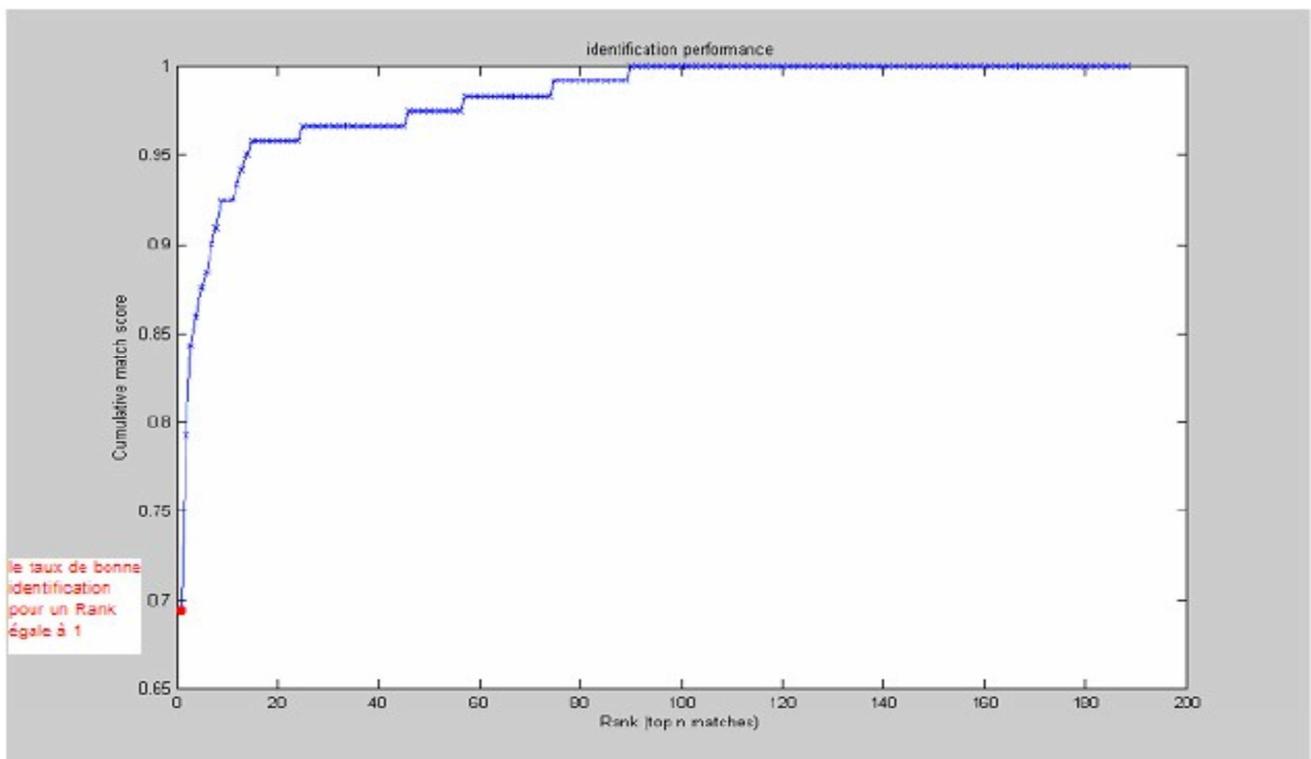
Dans la première phase de tests, l'objectif est de calculer le taux de bonne classification des personnes.

Après avoir effectué l'apprentissage, un système de nouvelles images est présenté (des images de test). La véritable question n'est pas toujours "est ce que l'identité la plus proche décidée par le système est correcte?" mais "est ce que la réponse correcte est parmi les k identités les plus proches?" [29].

Pour chaque image de test, le système va calculer les distances séparant cette image avec l'ensemble des images d'apprentissage.

Pour chaque rang (Rank), on calcule le taux d'identification : pour un Rank = k, on prend pour toute image de test, les k plus petites distances (les k plus proches images d'apprentissage). Si l'identité de l'image de test concorde avec au moins une des identités des k images d'apprentissage, alors on décide que cet image a été bien classifiée, sinon on décide qu'elle a été mal classifiée. Et on calcule finalement pour se rang le pourcentage de bonne identification portant sur l'ensemble des images de test.

On calcule le taux pour différentes valeurs du rang et on affiche ces taux-là, sous forme d'un graphe normalisé, qui a été adopté dans le protocole d'évaluation des algorithmes de reconnaissance des visages FERET [29].



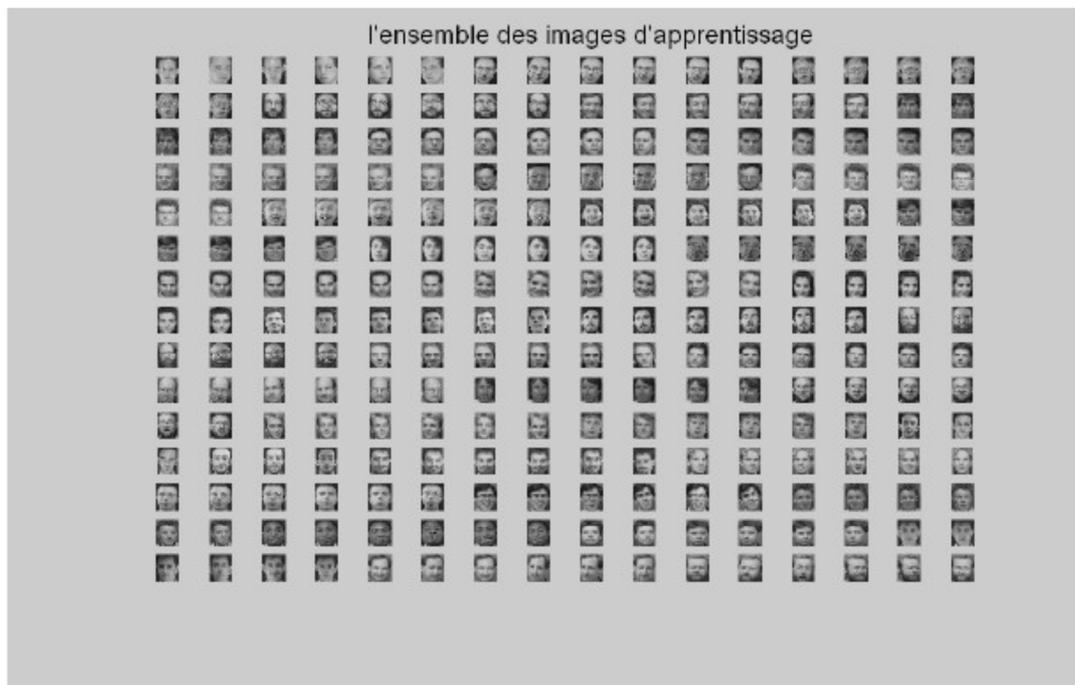
**FIG. III.1- Graphe des performances d'identification**

L'axe horizontal de ce graphe représente la valeur du Rank (on varie la valeur du Rank de 1 jusqu'au nombre d'images d'apprentissage). L'axe vertical quant à lui, représente le taux de bonne identification.

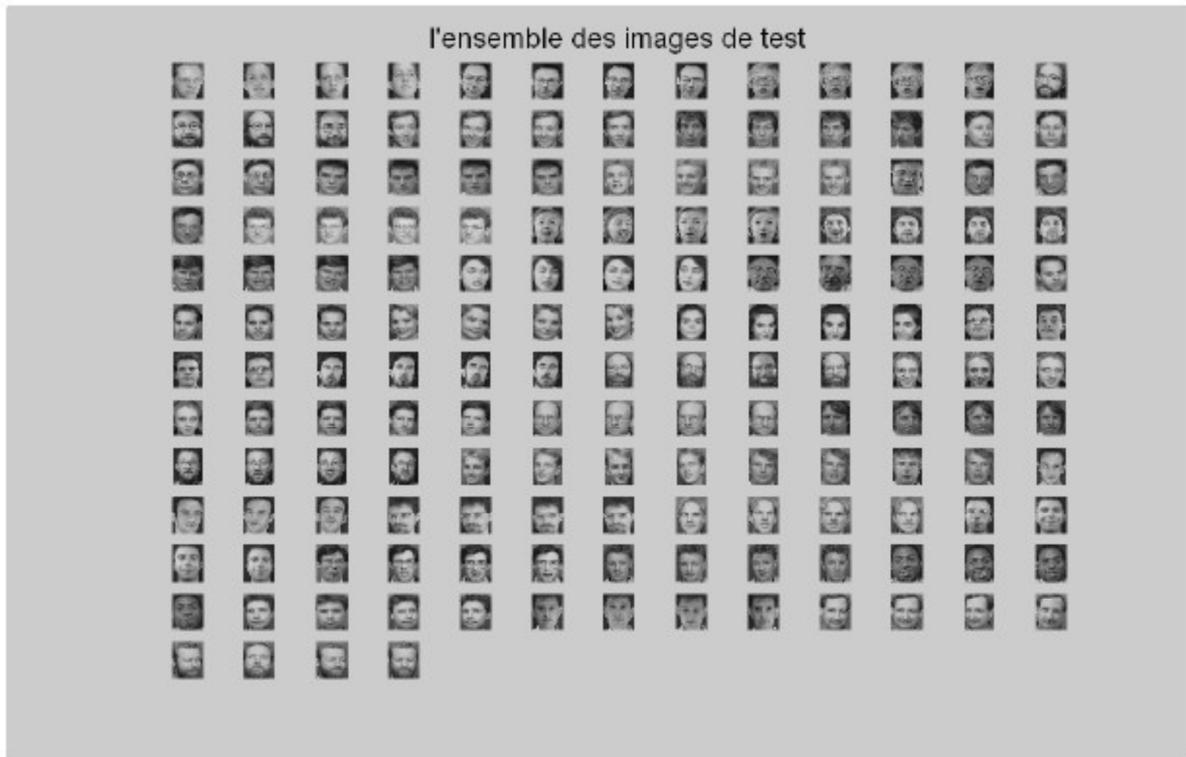
La comparaison des graphes relatifs aux différents algorithmes, nous permettra de déterminer l'algorithme qui se trompe le moins (ayant le plus grand taux pour un Rank égal à 1). Dans le cas de deux algorithmes avec un même taux pour un Rank égal à 1, le meilleur algorithme sera celui avec la plus grande variation du taux d'identification (la courbe ayant la plus grande pente).

### III.3.1 Approches Globales et Locales

Pour les premiers tests, la totalité des images de la base AT&T ont été utilisées. 240 images dans la phase d'apprentissage (40 personnes, avec 6 photos par personne), et 160 images de test à identifier (40 personnes, avec 4 photos par personne).



**FIG. III.2- L'ensemble des images d'apprentissage de la base AT&T utilisées pour le calcul du taux de la bonne identification**



***FIG. III.3- L'ensemble des images de test de la base AT&T utilisées pour le calcul du taux de la bonne identification***

Et voici les résultats de la bonne identification des algorithmes PCA, DF-LDA, LVQ, SVM, Template-matching, les CROSS RATIO et la distance de Hausdorff (les graphes des performances d'identification figurent dans l'annexe A) :

Les résultats obtenus pour les méthodes globales :

- PCA : pour un Rank égal à 1, on a 93.75% de bonne identification (FIG. A.1, annexe A).
- DF-LDA : pour un Rank égal à 1, on a 94.375% de bonne identification (FIG. A.2, annexe A).

➤ Le réseau de neurones LVQ :

L'apprentissage est fait en ajustant continuellement durant un certain nombre d'epochs les poids de connexion et les polarisations (biases), afin que le réseau apprenne à bien classifier les images d'apprentissage. Un cycle durant lequel les poids et les polarisations sont ajustés s'appelle l'epoch.

Lors de la phase d'apprentissage, le programme n'atteint jamais une erreur cumulative inférieure au seuil fixé (conditions d'arrêt si erreur =  $1e-10$ , ou soit 700 epochs) et il finit par s'arrêter après les 700 epochs (à peu près 1 heure d'exécution).

Le taux de bonne classification atteint 81.25%.

Ce taux peut relativement augmenter, en augmentant le nombre d'epochs, et en modifiant les paramètres du réseau.

Le réseau de neurones donne en sortie, la classe censée être la plus proche de l'image de test, mais par contre il ne permet pas de déterminer les autres classes qui suivent (2<sup>sd</sup> classe la plus proche ...), ainsi on obtient seulement le taux d'identification pour un rang égal à 1.

➤ SVM :

Un outil libre a été utilisé [30] [31] pour effectuer la reconnaissance de visages, en donnant comme entrée aux programmes d'apprentissage et de classification, des images de visages. On a adopté l'approche une contre toutes pour faire la classification multi classe, et on a fait une série de tests sur les kernels linéaires, polynomiaux (de degré 2, 3 et 4), sigmoïdes et gaussiens pour déterminer le kernel à utiliser, et finalement on a choisi le kernel polynomial de degré 3, vu qu'il offre les meilleurs performances.

Type du <i>kernel</i>	Linéaire	Polynomial de degré 2	Polynomial de degré 3	Polynomial de degré 4	sigmoïde	Gaussien
Performance d'identification	98.75%	96.25%	97.50%	97.50%	5.00%	96.25%
EER	16.1250%	17.5000%	<b>13.3750%</b>	13.6250%	56.8750%	21.3750%

***TAB. III.1- Les performances des kernels linéaires, polynomiaux, sigmoïdes et gaussiens sur AT&T***

Les performances de l'identification sont : pour un Rank égal à 1, on a 95% de bonne identification (FIG. A.3, annexe A).

- Template-matching : pour un Rank égal à 1, on a 86.25% de bonne identification (FIG. A.4, annexe A).

Les résultats obtenus pour les approches locales :

- Utilisation de la distance de Hausdorff, entre les positions des points caractéristiques des images AT&T (images de résolution 92 \* 112) : l'extraction des positions a été faite manuellement avec la souris, pour un Rank égal à 1, on a 24.37% de bonne identification (FIG. A.5, annexe A).
- Utilisation des CROSS RATIO entre les positions des points caractéristiques des images AT&T (images de résolution 92 \* 112) : l'extraction des positions a été faite manuellement avec la souris, pour un Rank égal à 1, on a 8.13% de bonne identification (FIG. A.6, annexe A).

La qualité de l'extraction des positions des points caractéristiques et la résolution des images utilisées ont influencé sur les performances des méthodes locales.

### III.3.2 Approches Hybrides

Au lieu d'appliquer l'algorithme PCA sur des images de visage, on appliquera PCA sur uniquement des régions des yeux et de la bouche, en donnant comme entrée à l'algorithme, un vecteur résultant de la concaténation des vecteurs de ces régions (on concatène les valeurs des pixels de l'image de l'œil gauche, avec les valeurs des pixels de l'image de l'œil droite, et avec celles de l'image de la bouche).

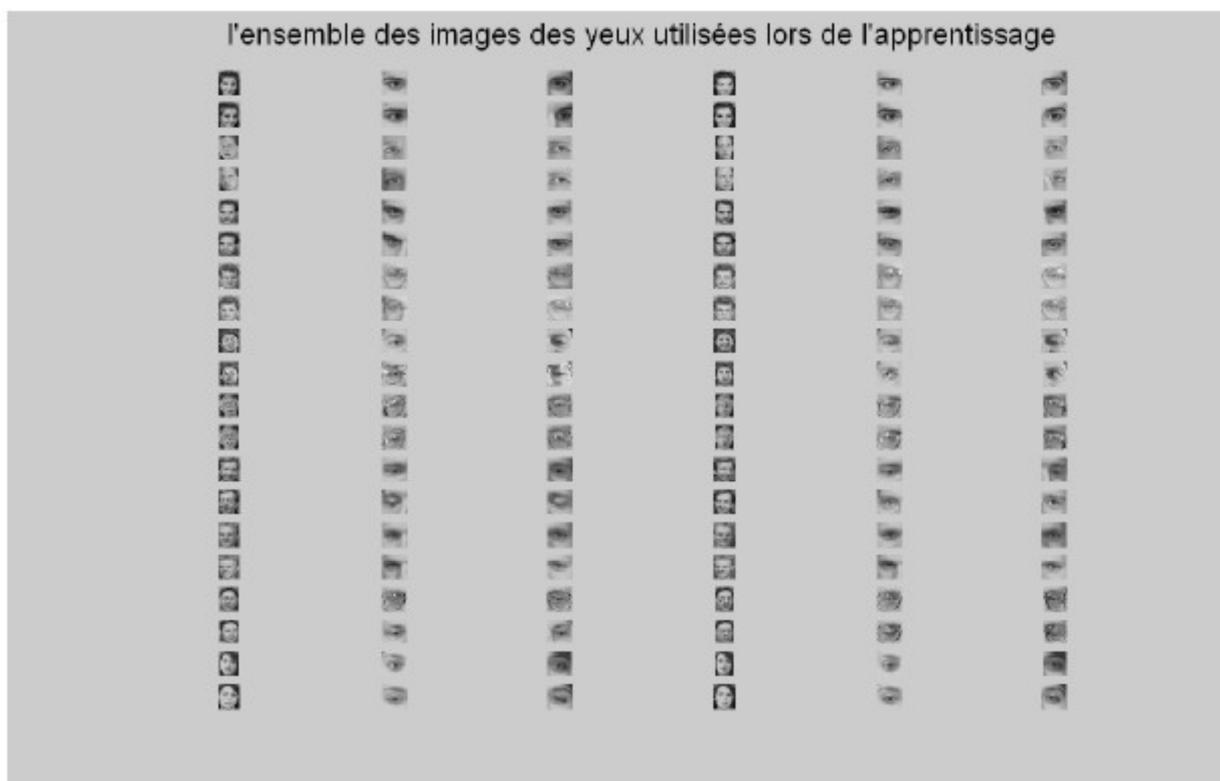
On a extrait ces régions, en utilisant "la fonction imcrop de Matlab", à partir des images de la base. On a utilisé uniquement les images convenables de la base AT&T (élimination des images avec des yeux fermés, et des images avec des grimaces touchant la région de la bouche). La phase d'apprentissage contient 188 images, et la phase de test 121 images à identifier.

Voici les résultats obtenus :

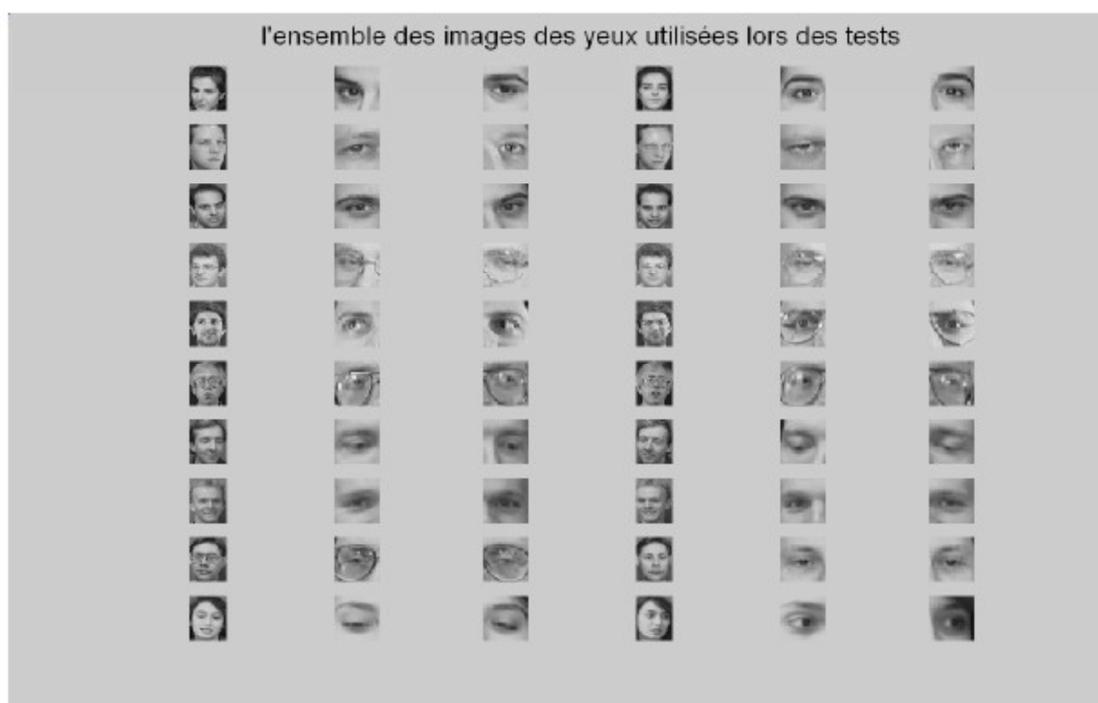
- Utilisation de l'œil gauche et de l'œil droit, comme entrée au processus PCA : pour un Rank égal à 1, on a 69.42% de bonne identification (FIG. A.7, annexe A).
  
- Utilisation de l'œil gauche et de l'œil droit et de la bouche, comme entrée au processus PCA : pour un Rank égal à 1, on a 80.17% de bonne identification (FIG. A.8, annexe A).

Pour valider les tests réalisés dans le cadre de la reconnaissance de visages en se basant sur l'extraction des points caractéristiques à savoir les régions des yeux et de la bouche, on a effectué des tests en utilisant la méthode LVQ sur uniquement les yeux au lieu du visage entier, ces images ont été extraites à partir des images de visages par les créateurs de la base AT&T.

On a pris 40 images d'apprentissage (4 images par personne) et 20 images de test à identifier (2 images par personne).



*FIG. III.4- L'ensemble des images des yeux de la base AT&T utilisées lors de l'apprentissage pour le calcul du taux de la bonne identification*



*FIG. III.5- L'ensemble des images des yeux de la base AT&T utilisées lors des tests pour le calcul du taux de la bonne identification*

Après la phase d'apprentissage (200 epochs), le programme donne un taux de bonne classification égal à 75%. Ce taux peut relativement augmenter, en augmentant le nombre d'epochs, et en modifiant les paramètres du réseau.

### III.3.3 Tableau récapitulatif

Le tableau suivant regroupe l'ensemble des performances d'identification pour les différents algorithmes.

Le nombre d'images d'apprentissage : 240 (40 personnes, avec 6 photos par personne). Le nombre d'images de test : 160 (40 personnes, avec 4 photos par personne).	
SVM	95%
DF_LDA	94.375%
PCA	93.75%
Utilisation de la totalité de l'image du visage comme entrée au réseau de neurones LVQ	81.25%
Template-matching	86.25%
Utilisation de la distance de Hausdorff, entre les positions des points caractéristiques des images	24.37%
Utilisation des CROSS RATIO entre les positions des points caractéristiques des images	8.13%
Le nombre d'images d'apprentissage : 188. Le nombre d'images de test : 121.	
Utilisation de l'œil gauche et de l'œil droit et de la bouche, comme entrée au processus PCA	80.17%
Utilisation de l'œil gauche et de l'œil droit, comme entrée au processus PCA	69.42%
Le nombre d'images d'apprentissage : (40 images, extraction AT&T, 4 images par personne). Le nombre d'images de test : (20 images, extraction AT&T, 2 images par personne).	
Utilisation de l'œil droite et de l'œil gauche comme entrée au réseau de neurones LVQ	75%

**TAB. III.2- Les performances d'identification des algorithmes SVM, DF-LDA, PCA, LVQ, Template-matching, la distance de Hausdorff, les CROSS RATIO, PCA appliquée aux régions des yeux, PCA appliquée aux régions des yeux et de la bouche et LVQ appliquée aux régions des yeux**

On remarque que les méthodes globales sont plus performantes que les méthodes locales et hybrides.

### **III.4 La méthode adoptée**

Les tests effectués auparavant, ont permis de conclure que PCA, LDA et SVM sont les méthodes qui donnent les meilleures performances.

Et pour choisir une méthode finale entre les trois, une étude comparative plus approfondie et des tests doivent être effectués afin d'en sortir avec une seule méthode.

#### **III.4.1 PCA Vs LDA**

##### **III.4.1.1 Les principales différences entre PCA et LDA**

Lors de tests [32] sur la base UMIST, on a utilisé 60 images d'apprentissage (3 images par personne : une vue frontale, latérale, et de 45-degree) et 20 images de test (20 personnes) avec des poses différentes. Les résultats obtenus montrent que :

- Lors de l'utilisation de LDA : 16 des 20 images ont été bien classifiées, en ayant souvent (13 fois) les 3 images les plus proches correctes.
- Lors de l'utilisation de PCA : 14 des 20 images ont été bien classifiées. Mais il faut noter que, le cas où les 3 images les plus proches sont tous correctes, n'a pas été rencontré.

Clairement on a constaté que LDA est plus performante sous la variation de pose, quand on dispose de peu d'image d'apprentissage sous différentes poses, ceci est dû au fait que la méthode LDA réduit au minimum l'éparpillement intra classe.

Le tableau suivant regroupe les principales différences entre LDA et PCA :

	LDA	PCA
Complexité informatique (calcul)	Plus complexe	Simple
Complexité de la méthode (implémentation)	Très complexe	Simple
Efficacité à travers la pose	Bonne même avec des données limitées	Peu, avec assez de données
Sensibilité au changement d'éclairage	Peu (utilisation avec divers illumination)	Très
Cas d'une image par classe	Moins performante	Plus performante
Taille des images (résolution)	Plus petite	Plus grande
Temps d'apprentissage	Beaucoup plus long (496.9370 seconde pour 240 images de la base AT&T)	Moins long (67.0150 seconde pour 240 images de la base AT&T)
Temps de reconnaissance	Inférieur à 0.5 seconde	Inférieur à 0.5 seconde

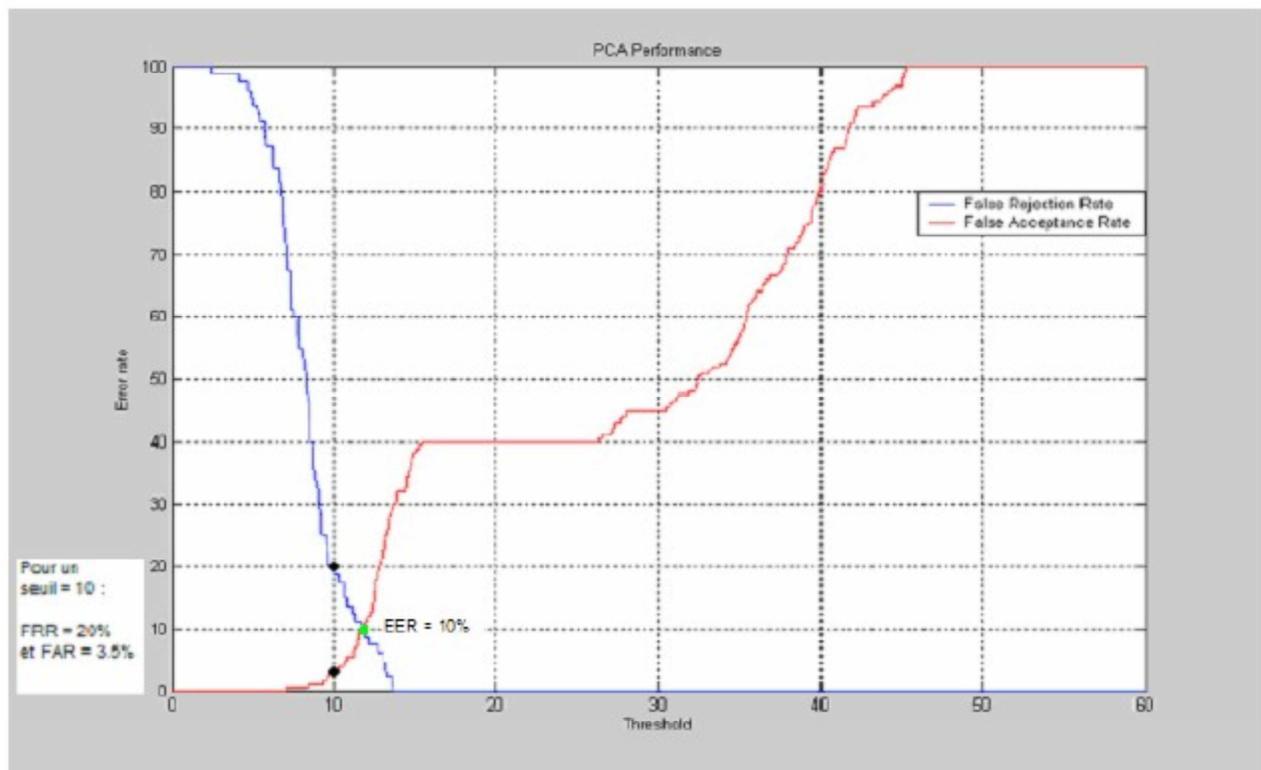
**TAB. III.3- Les principales différences entre LDA et PCA**

#### **III.4.1.2 L'utilisation d'un seuil pour l'élimination des personnes inconnues**

Lors de la phase de décision, tout système de reconnaissance associé à l'image à identifier, qu'elle soit d'une personne connue ou inconnue, une image (la plus proche) de la base d'apprentissage. D'où la nécessité d'utiliser un seuil de décision pour permettre l'accès seulement aux personnes connues du système.

La méthode de reconnaissance optimale est celle qui n'accepte pas les inconnus (taux de fausse acceptation  $FAR = 0$ ) et qui ne rejette aucun des connus (taux de faux rejet  $FRR = 0$ ). En réalité on ne peut pas avoir ceci, c'est pour cela qu'on cherche un seuil qui garantit à la fois un faible  $FRR$  et un faible  $FAR$ .

Pour trouver le seuil optimum, on calcule les  $FRR$  et  $FAR$  pour différentes valeurs du seuil. L'ensemble de ces résultats est représenté sous forme d'un graphe standard (courbe ROC [33]), qui est utilisé lors de la comparaison entre les algorithmes.



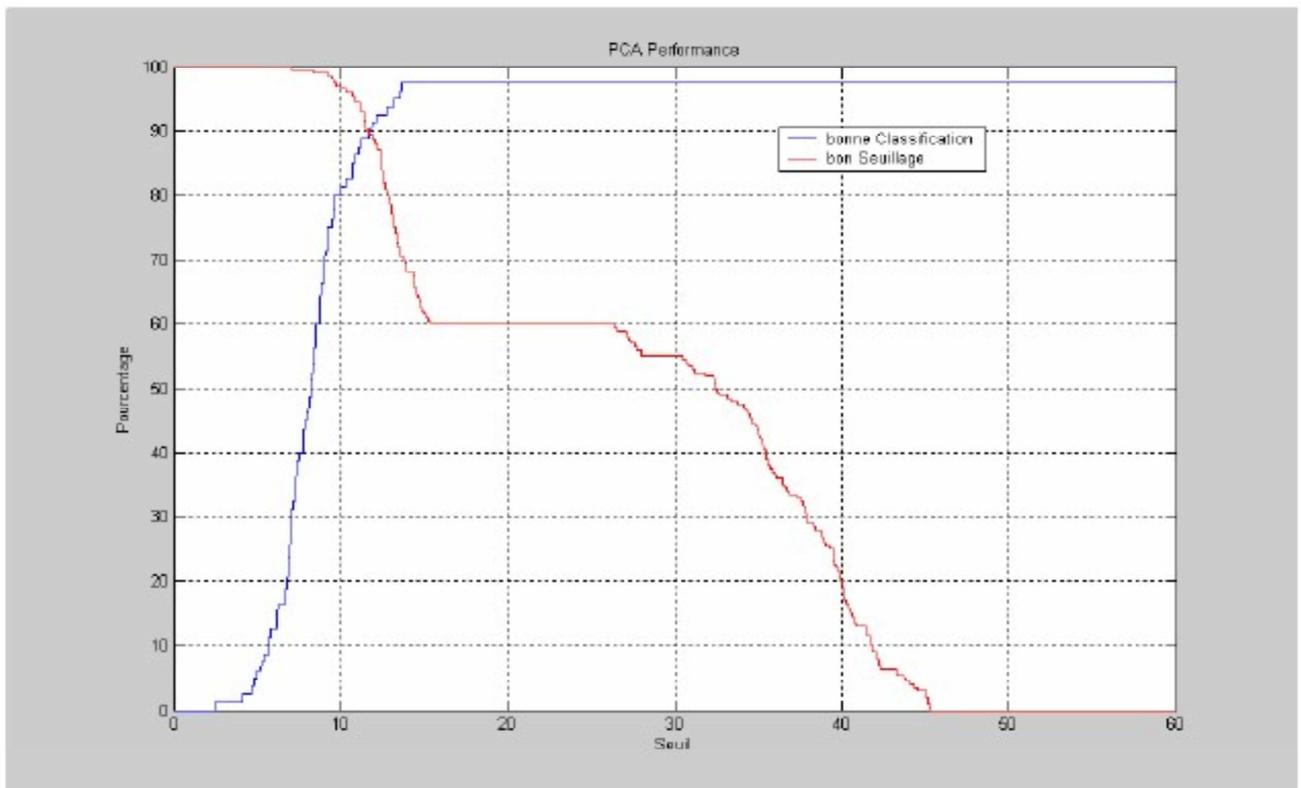
**FIG. III.6- Graphe des valeurs de  $FRR$  et de  $FAR$  en fonction de la valeur du seuil**

L'axe horizontal de ce graphe représente la valeur du seuil et l'axe vertical représente les valeurs de  $FRR$  et de  $FAR$ .

Toute base de données à utiliser, est divisée en deux parties : les employés et les inconnus. Les images des employés sont elles aussi réparties en deux catégories : les images d'apprentissage et les images de test (à vérifier l'identité).

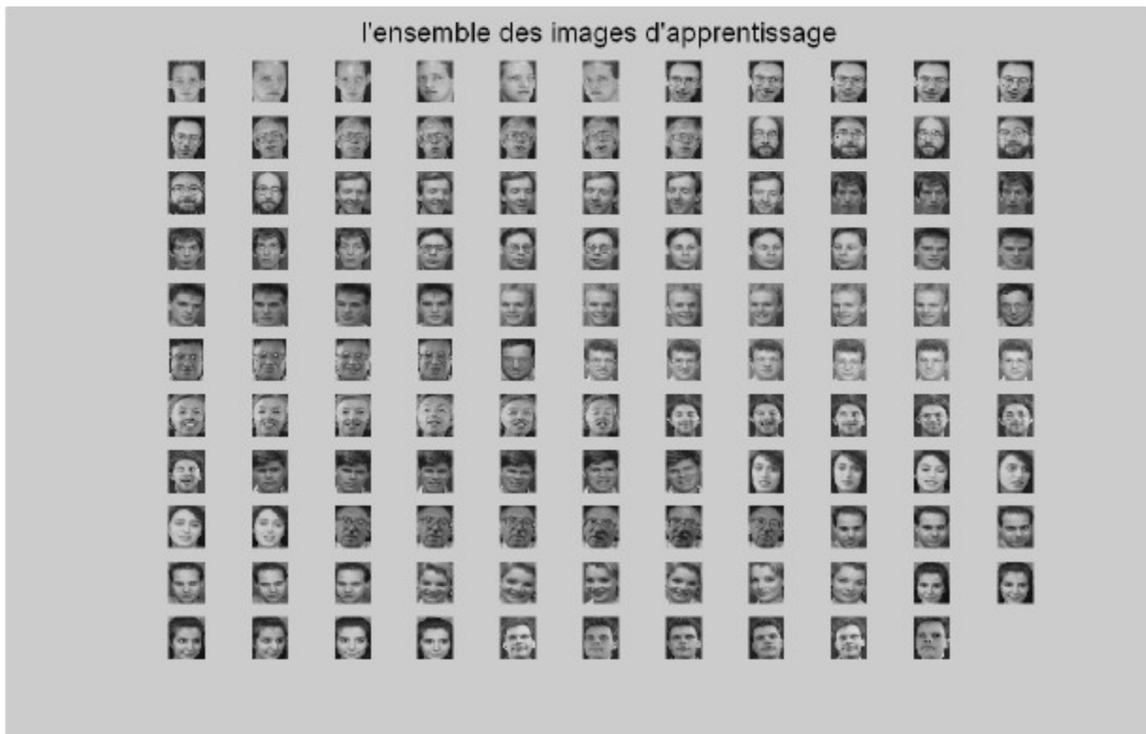
Dans la courbe du FAR et du FRR en fonction du seuil, on a l'information du taux du non rejet des sujets connus (acceptés par le système). C'est sujets là, peuvent être soit bien identifiés soit mal identifiés par le système.

Et pour avoir l'information du taux de la bonne identification, on trace une courbe qui donne le taux de la bonne élimination (bon seuillage) des inconnus et le taux de la bonne classification des sujets connus, en fonction de la valeur du seuil.

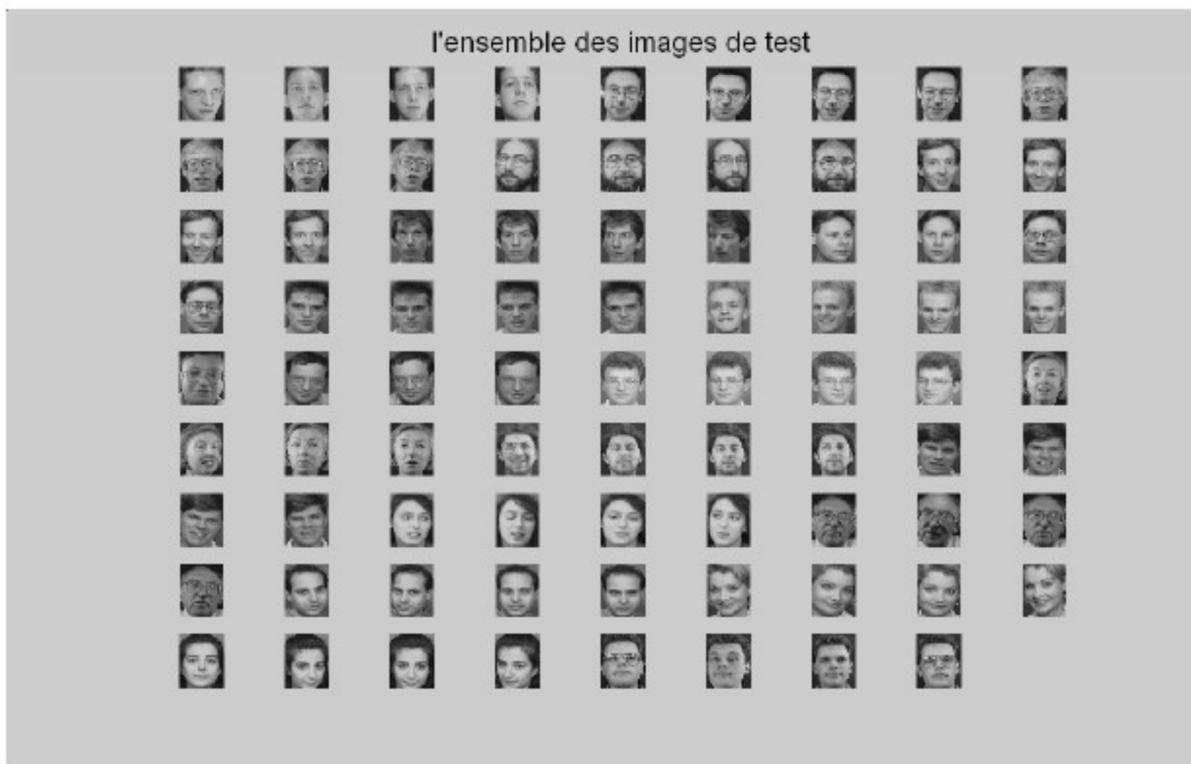


**FIG. III.7- Graphe des taux du bon seuillage des inconnus et des taux de la bonne classification des sujets connus en fonction de la valeur du seuil**

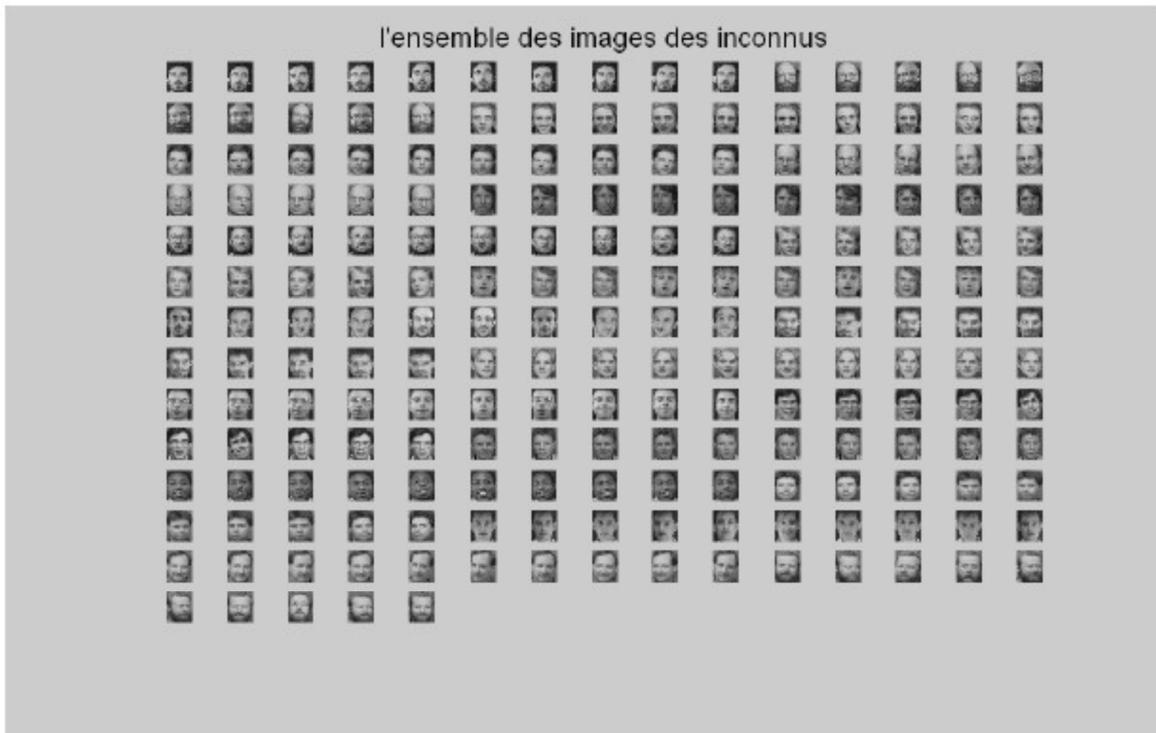
Pour les tests on a utilisé la base AT&T. On a pris 120 images d'apprentissage (20 personnes, avec 6 photos par personne), 80 images de test (20 personnes, avec 4 photos par personne), et 200 images d'inconnus (20 personnes, avec 10 photos par personne).



*FIG. III.8- L'ensemble des images d'apprentissage de la base AT&T utilisées pour la détermination d'un seuil pour l'élimination des personnes inconnues*



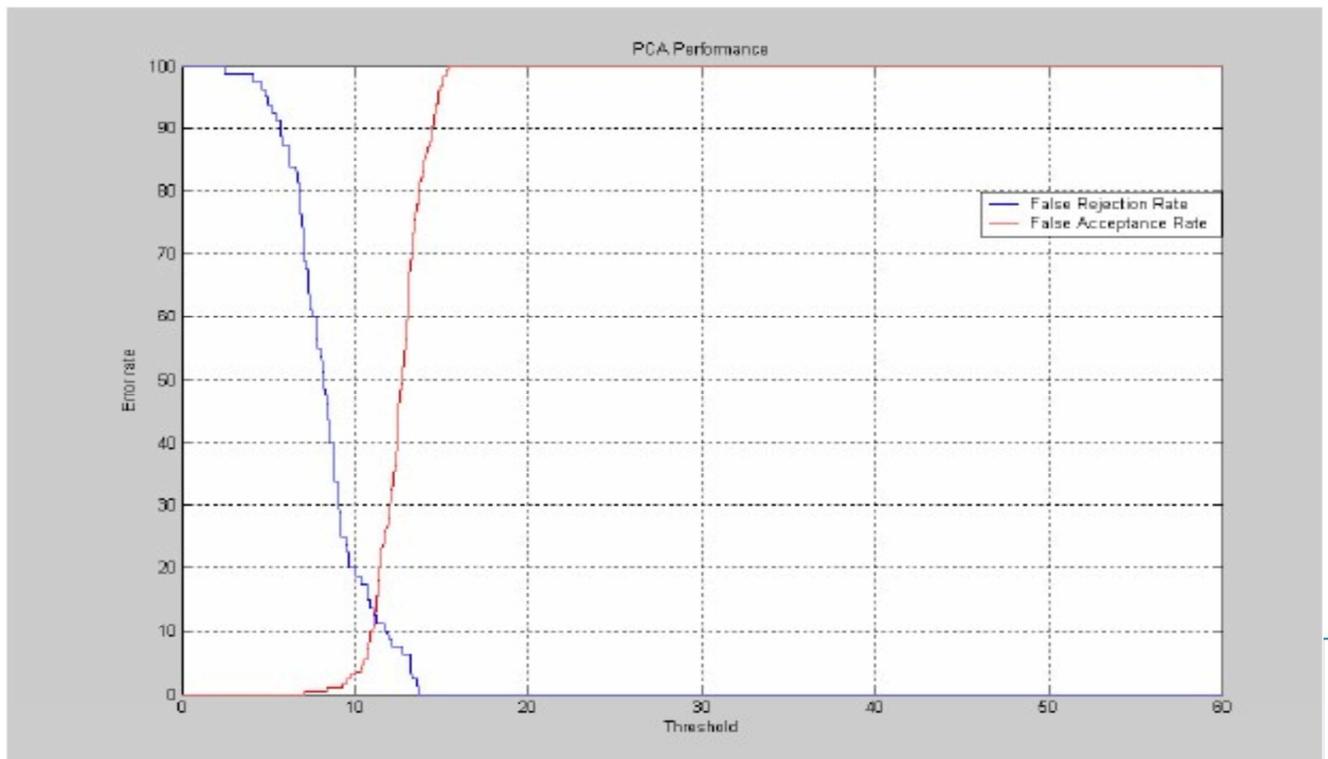
**FIG. III.9-** L'ensemble des images de test de la base AT&T utilisées pour la détermination d'un seuil pour l'élimination des personnes inconnues



**FIG. III.10-** L'ensemble des images des inconnus de la base AT&T utilisées pour la détermination d'un seuil pour l'élimination des personnes inconnues

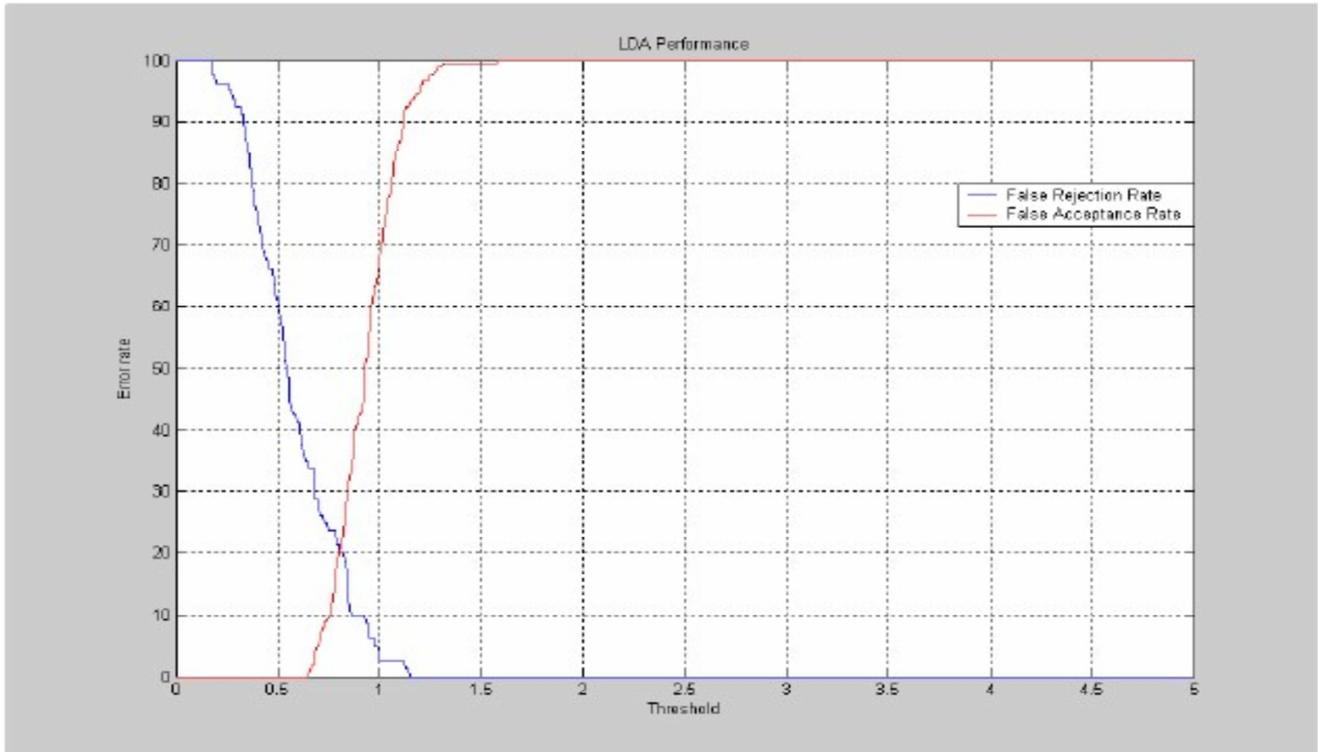
Et voici les graphes relatifs aux deux algorithmes testés : PCA et LDA.

➤ PCA :



**FIG. III.11-** Le graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil, de PCA

➤ LDA :



**FIG. III.12-** Le graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil, de LDA

On constate que lors de l'utilisation d'un seuil constant pour l'élimination des personnes inconnues, PCA est plus performante (elle garantit un bon compromis offrant à la fois un bas FAR et un bas FRR).

La méthode LDA est à utiliser dans le cas d'une application où on a des problèmes de pose ou d'éclairage. Mais, si on peut déterminer et avoir les meilleures conditions de pose et d'éclairage, la méthode PCA constituera alors la méthode la plus adéquate offrant à la fois de la rapidité, de la simplicité, des performances de classification aussi meilleures que celles de LDA, et surtout un pourcentage de rejet de personnes inconnues qui est élevé.

### III.4.2 PCA Vs SVM

On a testé PCA, SVM et des algorithmes combinés basés sur SVM, à savoir : PCA + SVM, LDA + SVM et ICA + SVM, sur 3 bases de données. Dans ces algorithmes combinés, l'entrée de SVM est le vecteur de poids (la projection) de l'image, résultant de l'application d'une méthode de réduction d'espace (PCA, LDA ou ICA).

On a utilisé dans ces tests l'algorithme d'ICA : FastICA [34], pour récupérer les vecteurs de poids ICA des images de visages.

Et l'ensemble des résultats obtenus est :

➤ Utilisation de la base AT&T :

120 images d'apprentissage (20 personnes, 6 photos par personne), 80 images de test (20 personnes, 4 photos par personne), et 200 images représentant des inconnus (20 personnes, 10 photos par personne).

	PCA	SVM	PCA+SVM	LDA+SVM	ICA+SVM
Bonne identification (apprentissage)	100%	100%	100%	100%	100%
Bonne identification (test)	97.50%	97.50%	97.50%	93.75%	96.25%
Contraintes d'utilisation	PCA sans phase de pré traitement		PCA avec phase de pré traitement		Vecteur de poids ICA de taille 100

**TAB. III.4- Les résultats des performances d'identification de PCA, de SVM et des algorithmes combinés basés sur SVM, sur AT&T**

La première ligne du tableau donne les taux de la bonne identification des images

d'apprentissage, la deuxième ligne donne les taux de la bonne identification des images de test et la troisième ligne présente les contraintes d'utilisation des algorithmes, qui donnent les meilleures performances (la phase de pré traitement est composée d'une égalisation d'histogramme).

	PCA	SVM	PCA+SVM	LDA+SVM	ICA+SVM
EER	12.5000%	12.7500%	22.2500%	29.7500%	17.5000%
Pour 100% de bon seuillage	42.5000% de bonne classification	40.0000% de bonne classification	12.5000% de bonne classification	10.0000% de bonne classification	56.2500% de bonne classification

**TAB. III.5- Les résultats des performances de seuillage (des personnes inconnues) de PCA, de SVM et des algorithmes combinés basés sur SVM, sur AT&T**

La première ligne du tableau donne les valeurs du EER et la deuxième ligne donne le pourcentage de la bonne classification des personnes connues, dans le cas de l'utilisation d'un seuil qui bloque la totalité des personnes inconnues.

➤ Utilisation de la base UMIST :

165 images d'apprentissage (11 personnes, 15 photos par personne), 418 images de test (11 personnes, 38 photos par personne), et 302 images représentant des inconnus.

	PCA	SVM	PCA+SVM	LDA+SVM	ICA+SVM
Bonne identification (apprentissage)	100%	100%	98.18%	93.33%	100%
Bonne identification (test)	96.17%	96.41%	92.11%	88.52%	96.86%
Les contraintes d'utilisation	PCA sans phase de pré traitement		PCA sans phase de pré traitement		Vecteur de poids ICA de taille 100

**TAB. III.6- Les résultats des performances d'identification de PCA, de SVM et des algorithmes combinés basés sur SVM, sur UMIST**

	PCA	SVM	PCA+SVM	LDA+SVM	ICA+SVM
EER	8.3257%	7.9209%	26.1177%	53.6155%	12.6311%
Pour 100% de bon seuillage	<b>89.2344% de bonne classification</b>	72.9665% de bonne classification	37.3206% de bonne classification	2.3923% de bonne classification	28.2297% de bonne classification

**TAB. III.7- Les résultats des performances de seuillage (des personnes inconnues) de PCA, de SVM et des algorithmes combinés basés sur SVM, sur UMIST**

➤ Utilisation de la base YALE :

70 images d'apprentissage (10 personnes, 7 photos par personne), 40 images de test (10 personnes, 4 photos par personne), et 55 images représentant des inconnus (5 personnes, 11 photos par personne).

	PCA	SVM	PCA+SVM	LDA+SVM	ICA+SVM
Bonne identification (apprentissage)	100%	98.57%	100%	84.29%	100%
Bonne identification (test)	90.00%	82.50%	97.50%	77.50%	100%
Les contraintes d'utilisation	PCA avec phase de pré traitement		PCA avec phase de pré traitement		Vecteur de poids ICA de taille 100

**TAB. III.8- Les résultats des performances d'identification de PCA, de SVM et des algorithmes combinés basés sur SVM, sur YALE**

	PCA	SVM	PCA+SVM	LDA+SVM	ICA+SVM
EER	25.2273%	27.3864%	23.0683%	34.7727%	9.5455%
Pour 100%	27.5000% de	10.0000% de	22.5000% de	2.5000% de	25.0000% de

de bon seuillage	bonne classification				
------------------	----------------------	----------------------	----------------------	----------------------	----------------------

**TAB. III.9- Les résultats des performances de seuillage (des personnes inconnues) de PCA, de SVM et des algorithmes combinés basés sur SVM, sur YALE**

On veut que notre système n'accepte aucun inconnu, donc on conclut d'après ces résultats que dans le cas d'application de PCA, on a les meilleures performances (le plus haut taux de la bonne classification pour un rejet absolu de tous les inconnus).

Donc la méthode qu'on a choisie pour le système final est PCA, et l'étape suivante consiste à améliorer ses performances.

### **III.5 L'amélioration des performances de PCA**

Dans la littérature on parle toujours de l'effet néfaste à la fois de la pose et de l'éclairage, qui détériorent les performances de tout système biométrique basé sur le visage. Et dans notre approche générale, visant à éviter le maximum d'étapes de pré traitement qui vont alourdir encore plus le système, on s'est intéressé à chercher les conditions optimales de pose et d'éclairage à considérer pour améliorer les performances et exploiter ainsi toute la puissance de l'algorithme.

#### **III.5.1 Résoudre le problème d'éclairage**

Le problème d'éclairage est généralement difficile à traiter et il a suscité beaucoup d'attentions. Pour la reconnaissance de visages, les chercheurs ont proposé beaucoup de bonnes approches, qu'on peut diviser en quatre types [35] :

➤ **Les approches heuristiques**

Dans le domaine du sous espace propre (eigen-subspace domain), il a été suggérer et vérifier, qu'en écartant les trois composantes principales les plus significatives, les variations dues à l'éclairage peuvent être réduites. Une autre méthode heuristique basée sur la symétrie de visage a été proposée pour augmenter les performances des systèmes sous différents éclairages.

- Les approches de comparaison d'images

Les approches statistiques basées sur la comparaison d'images utilisent différentes représentations d'images et mesures de distance. Parmi les représentations d'images employées on trouve : les cartes de contour, les dérivatives du niveau de gris, les images filtrées avec 2D Gabor-like functions, et les représentations combinant la fonction Log des intensités avec les autres représentations.

Différentes mesures de distance sont utilisées comme : point-wise distance, regional distance, affine-GL (gray level) distance, local affine-GL distance, et LOG point-wise distance. Il est à noter que ces représentations ne sont pas suffisantes par elles-mêmes pour surmonter les variations d'image.

- Les approches basées sur la classe

On utilise dans ces méthodes pour chaque visage, différentes images sous une pose fixe, mais avec différentes conditions d'éclairage.

- Les approches basées sur le modèle

Les méthodes basées sur le modèle emploient des modèles 3D.

### **III.5.2 Résoudre le problème de pose**

On a proposé diverses méthodes pour résoudre le problème de rotation. Ces méthodes peuvent être divisées en trois classes [35] :

- Les approches basées sur de multiples images

On emploie dans ces méthodes une multitude d'images sous différentes poses pour chaque personne.

- Les approches hybrides

De nombreux algorithmes ont été proposés dans ce type, qui reste le plus populaire. Et ceci est dû au fait qu'elles sont probablement les méthodes les plus réussies et les plus pratiques jusqu'à maintenant, en plus de leur utilisation de l'information préalable de la classe.

Parmi les méthodes les plus représentatives on trouve : linear class based method, graph matching based method, et view-based eigenface approach. En dépit de leur popularité, ces méthodes ont quelques inconvénients communs : le besoin de beaucoup d'images d'apprentissage pour couvrir toutes les vues possibles, et la séparation entre les deux problèmes de pose et d'éclairage.

- Les méthodes basées sur une unique image/forme

Et finalement, il y a la troisième classe des approches, qui incluent : *les low-level feature based methods, les invariant feature based methods, et la 3D model based method*. Cette troisième approche ne semble pas avoir suscité beaucoup d'attention.

### **III.5.3 Les conditions optimales de pose et d'éclairage**

En générale, il semble que tous les systèmes de reconnaissance basés sur le visage, adoptent tous les mêmes conditions.

Par exemple dans le système FaceVACS-SDK Version 4.0.3 qui a été développé et commercialisé par la grande firme allemande Cognitec Systems GmbH, on a fixé les règles suivantes [36] :

- la procédure d'apprentissage :

Une bonne procédure d'apprentissage, où on fait l'acquisition des images utilisées lors de cette phase, augmente sensiblement les performances de reconnaissance. Voici quelques procédés à suivre pour augmenter la qualité de l'apprentissage :

- Employez plusieurs images du visage, montrant des vues légèrement différentes. Ceci peut être accompli en demandant à l'utilisateur de faire de légers mouvements du visage dans la direction horizontale et verticale, ou en employant un arrangement approprié de plusieurs

appareils photo.

- Informez l'utilisateur que prendre des photos d'apprentissage n'est pas identique à faire un portrait chez le photographe. Ainsi il ne devrait ni poser ni sourire, mais montrer une expression journalière du visage, qui sera susceptible d'être la même lors de la vérification.

- Si le système éprouve des difficultés à reconnaître les personnes portant des lunettes, fusionnez des images avec et sans lunettes dans la phase d'apprentissage, ou demandez à la personne de faire l'apprentissage et la reconnaissance sans lunettes.

- l'influence et l'arrangement des conditions d'éclairage :

Dans les circonstances où on peut contrôler les conditions d'éclairage, on devrait prêter l'attention aux recommandations suivantes :

- L'éclairage du visage devrait être une lumière frontale diffuse ou dirigée, pour éviter les ombres volatiles dans la région de visage.

- Eviter l'éclairage produisant de la lueur dans les lunettes, ou sur une peau brillante.

- L'éclairage du visage ne devrait pas changer avec la position ou la taille de l'utilisateur.

- Protégez la zone biométrique contre l'éclairage additionnel, particulièrement contre, la lumière du soleil.

- Les conditions d'éclairage devraient être tout à fait semblables si non identique, à la fois dans la phase d'apprentissage et la phase de reconnaissance.

Dans beaucoup de cas, des tubes fluorescents placés à gauche et à droite de l'appareil photo, fournissent une illumination appropriée. Et s'ils sont placés décentrés à un degré d'approximativement 30-45 degrés, ça va éviter de causer de la lueur dans les lunettes.

### **III.5.4 Vérification**

Lors de tests réalisés sur 6 bases de données (AT&T, UMIST, HARVARD, FERET, YALE et une base comportant des images de visage de face) il a été prouvé que dans le cas où on a des images des personnes à identifier, dans chaque vue ou position possible et avec des conditions d'éclairage uniforme, ceci donne les performances optimales.

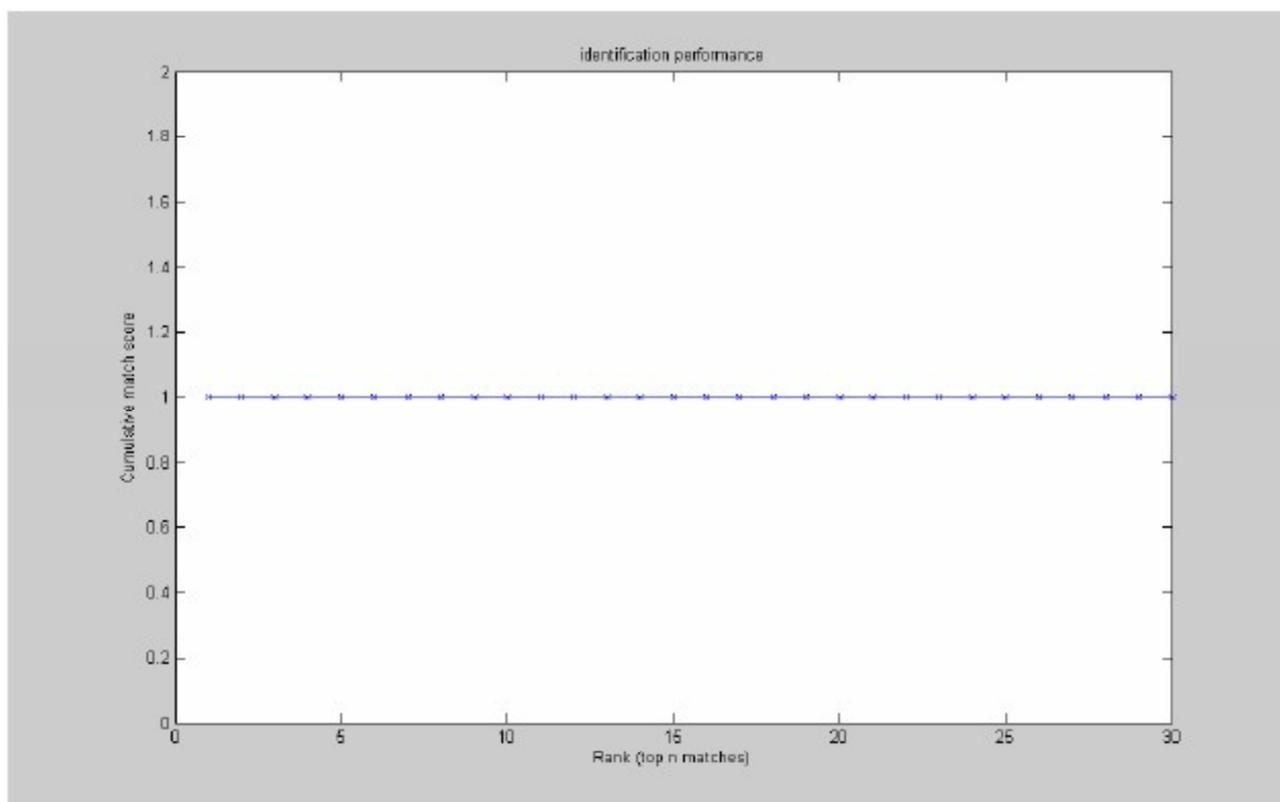
Les résultats de l'ensemble des tests réalisés sont :

➤ Utilisation d'une base comportant des images de visage de face :

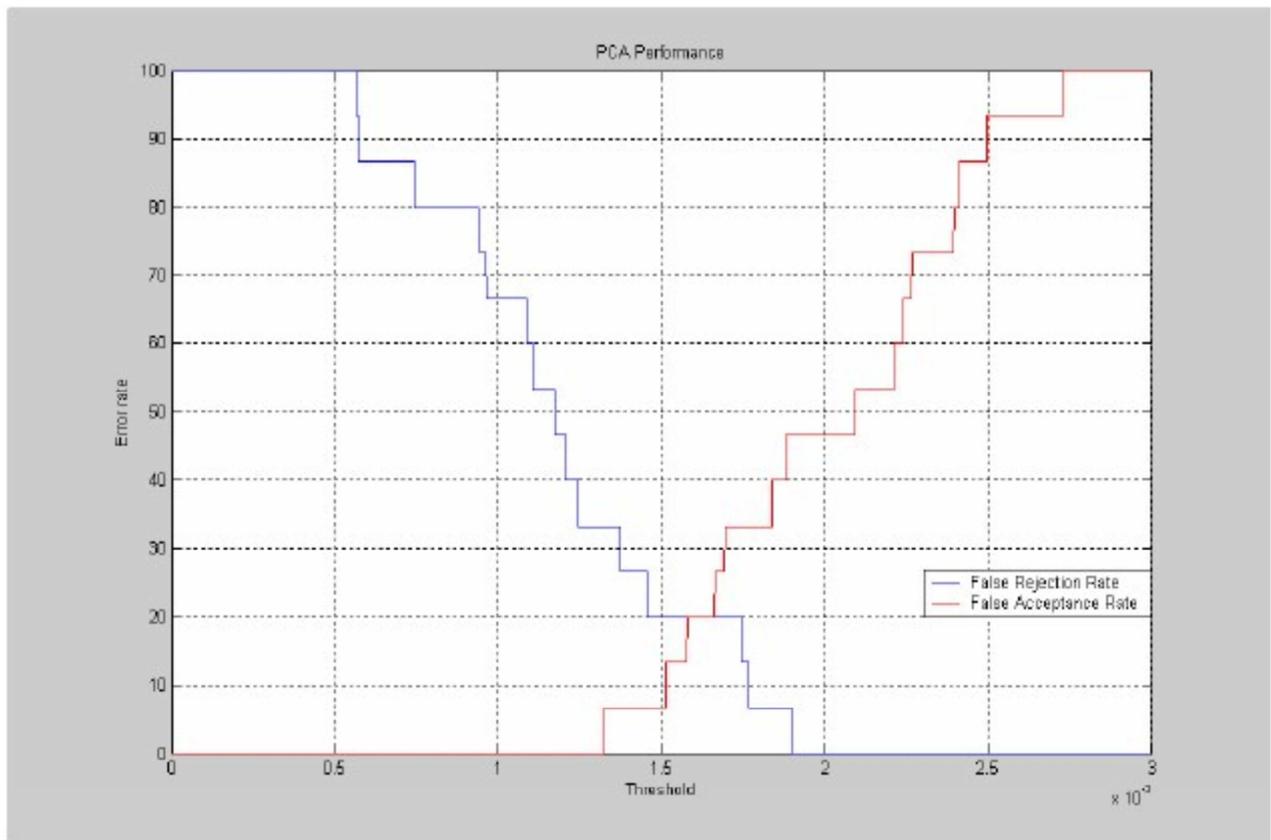
30 images de face d'apprentissage (15 personnes, 2 photos par personne), 15 images de face de test (15 personnes, 1 photos par personne), et 15 images de face représentant des inconnus (5 personnes, 3 photos par personne).

Les images utilisées sont véritablement en face (profile frontale), mais les conditions d'éclairage différent.

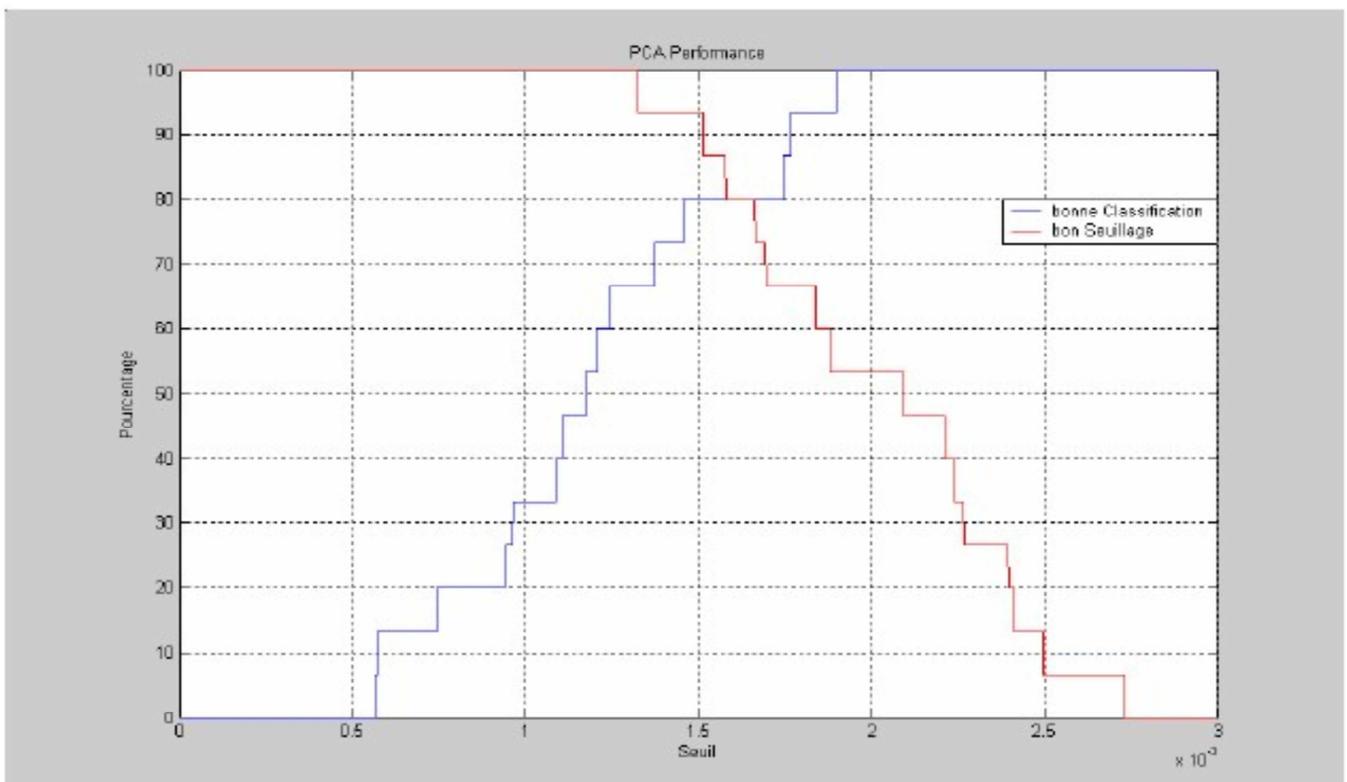
Le fait d'appliquer une phase de pré traitement (une égalisation d'histogramme) aux différentes images améliore les performances :



**FIG. III.13-** Le graphe des performances d'identification de PCA sur la base d'images de face



**FIG. III.14-** Le graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil, de PCA sur la base d'images de face



**FIG. III.15- Le graphe des taux du bon seuillage et de la bonne classification en fonction de la valeur du seuil, de PCA sur la base d'images de face**

➤ Utilisation de la base AT&T :

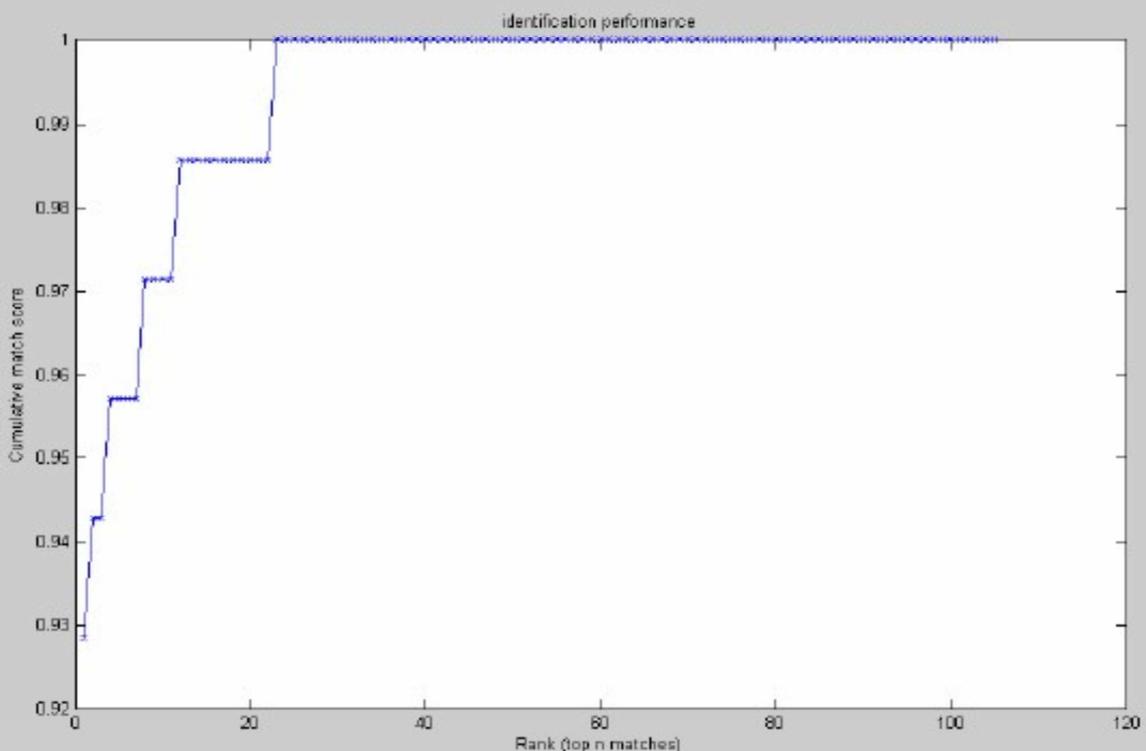
105 images de face, 69 images de face de test, et 71 images de face représentant les inconnus ont été prise Lors de l'apprentissage.

Les images utilisées ne sont pas bien positionnées (les personnes sont mal positionnées face à la camera, on n'a pas la même pose et posture à chaque fois).

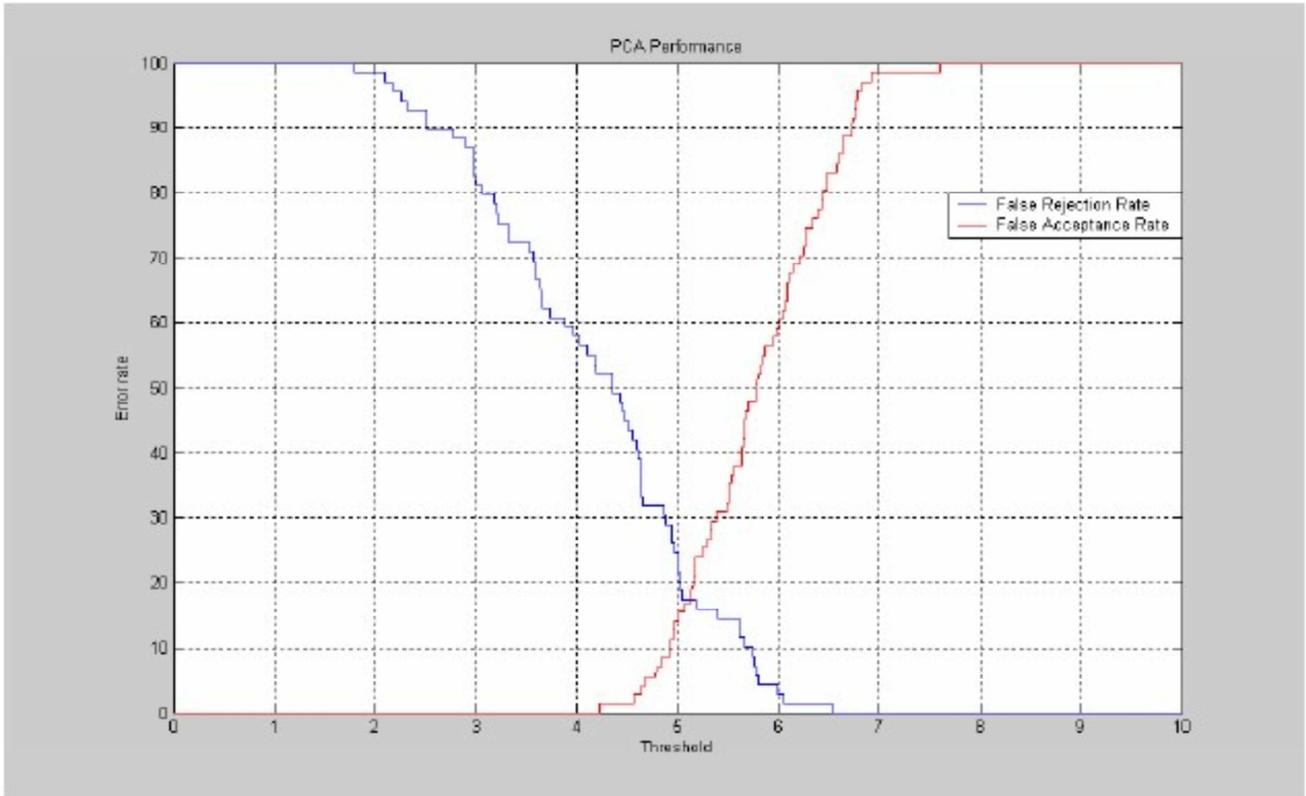
On a utilisé des images de personnes avec lunettes et / ou avec composantes structurales (barbres, moustaches, ... etc.). Les images utilisées résultent de l'application d'un cropage (la fonction `imcrop` de Matlab) suivie par une normalisation de la taille (la fonction `imresize` de Matlab) des images originales de la base.

L'application de la phase de pré traitement détériore visiblement les performances, c'est éventuellement dû à l'utilisation de `imcrop` et `imresize` (la perte de résolution résultante de `imcrop`, le changement et la modification des pixels produite par `imresize` ...).

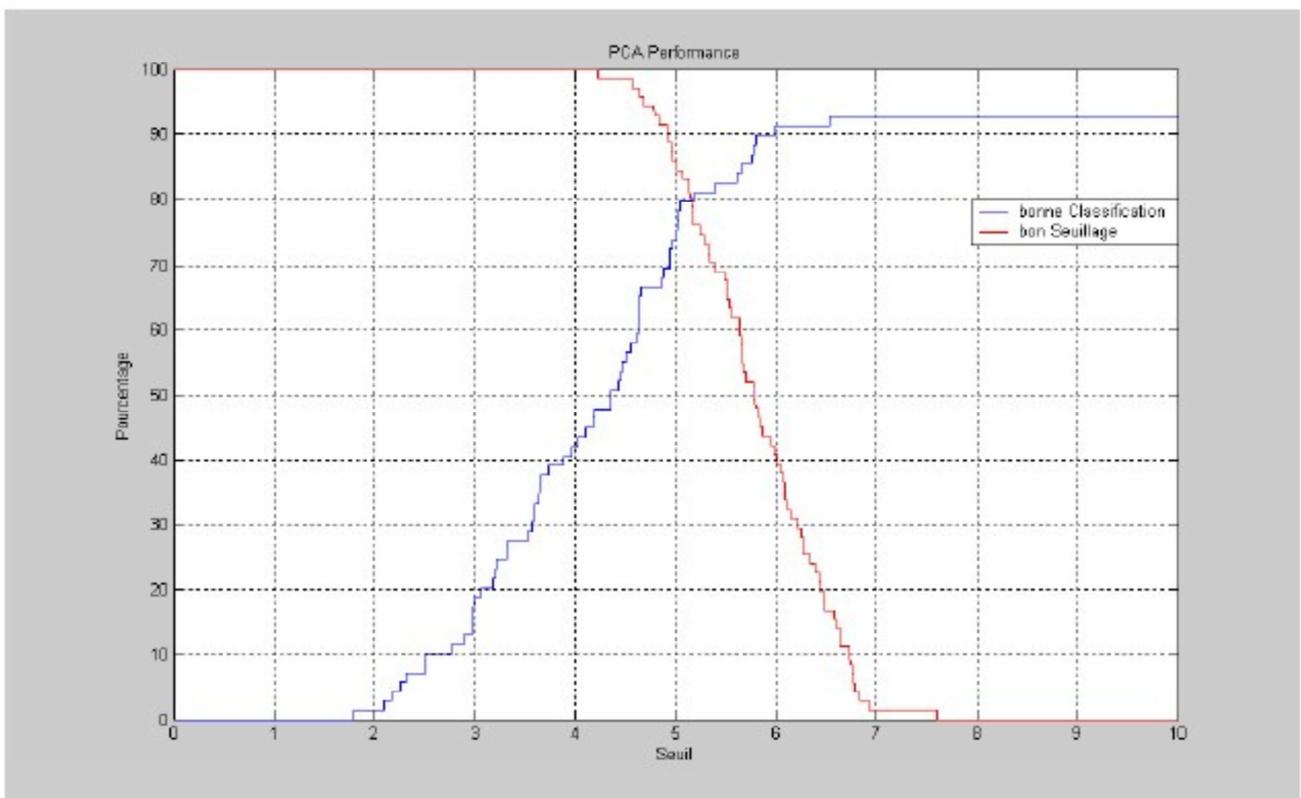
Et voici les performances d'identification et de seuillage :



**FIG.III.16- Le graphe des performances d'identification de PCA sur AT&T**



**FIG. III.17- Le graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil, de PCA sur AT&T**



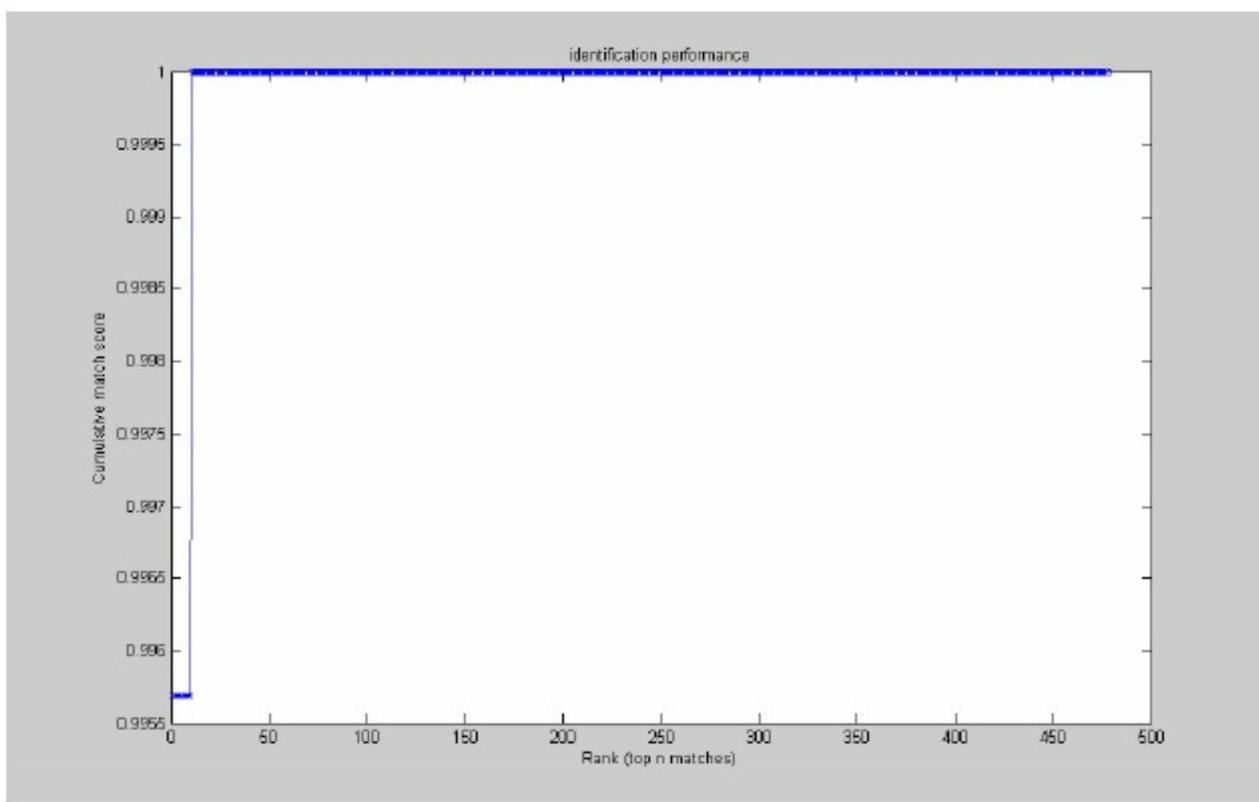
**FIG. III.18- Le graphe des taux du bon seuillage et de la bonne classification en fonction de la valeur du seuil, de PCA sur AT&T**

➤ Utilisation de la base UMIST :

478 images d'apprentissage, 232 images de test, et 302 images représentant les inconnus ont été prise.

La phase de pré traitement détériore les performances, (on a le même taux de bonne classification pour un Rank = 1, mais on atteint le 100% plus lentement), et c'est probablement dû à la normalisation et la réduction de la taille des images, qui avaient au départ différentes tailles.

Voici les performances de la bonne identification :



**FIG. III.19- Le graphe des performances d'identification de PCA sur UMIST**

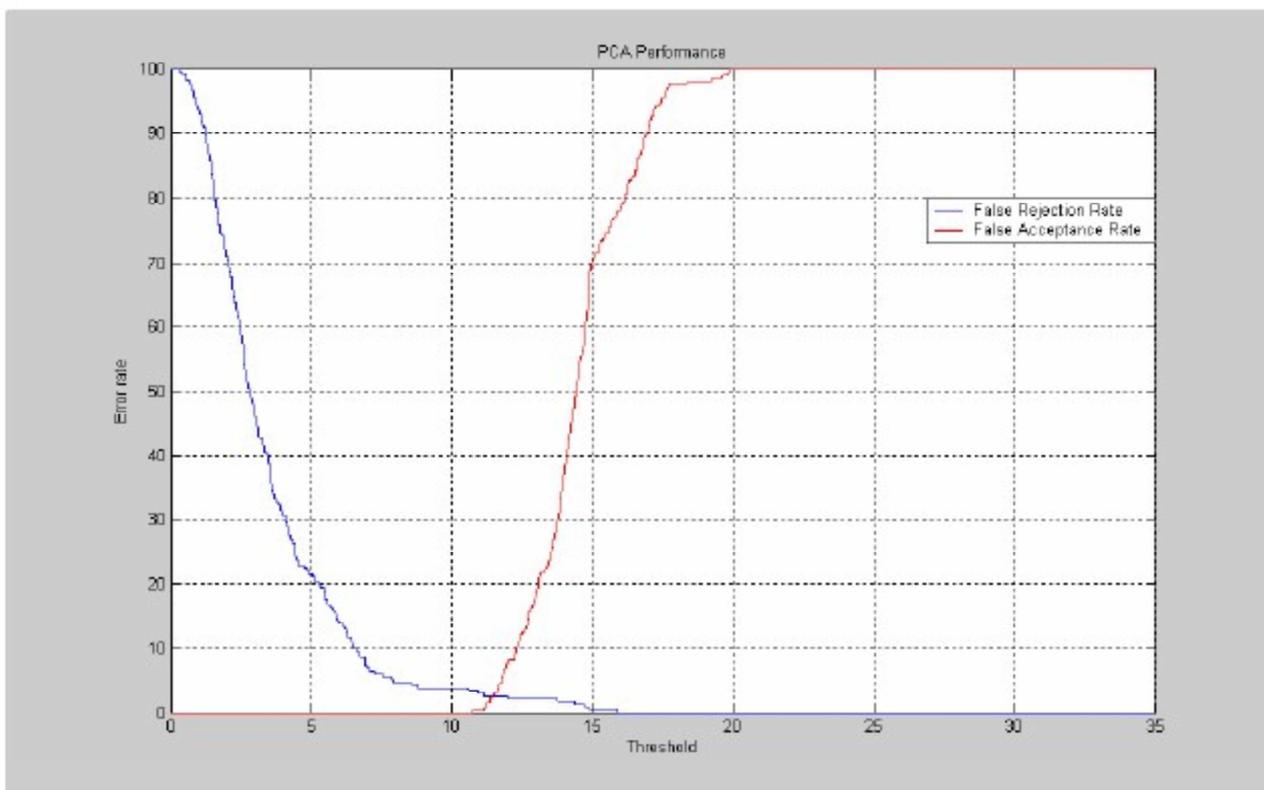
Aucune image de vérification n'a été utilisée dans la phase d'apprentissage. En réalité ce qui

explique ce taux de classification très élevé, et le fait que pour chaque image de vérification, il y'a une image d'apprentissage qui lui est proche (il y'a juste une petite variation de pose).

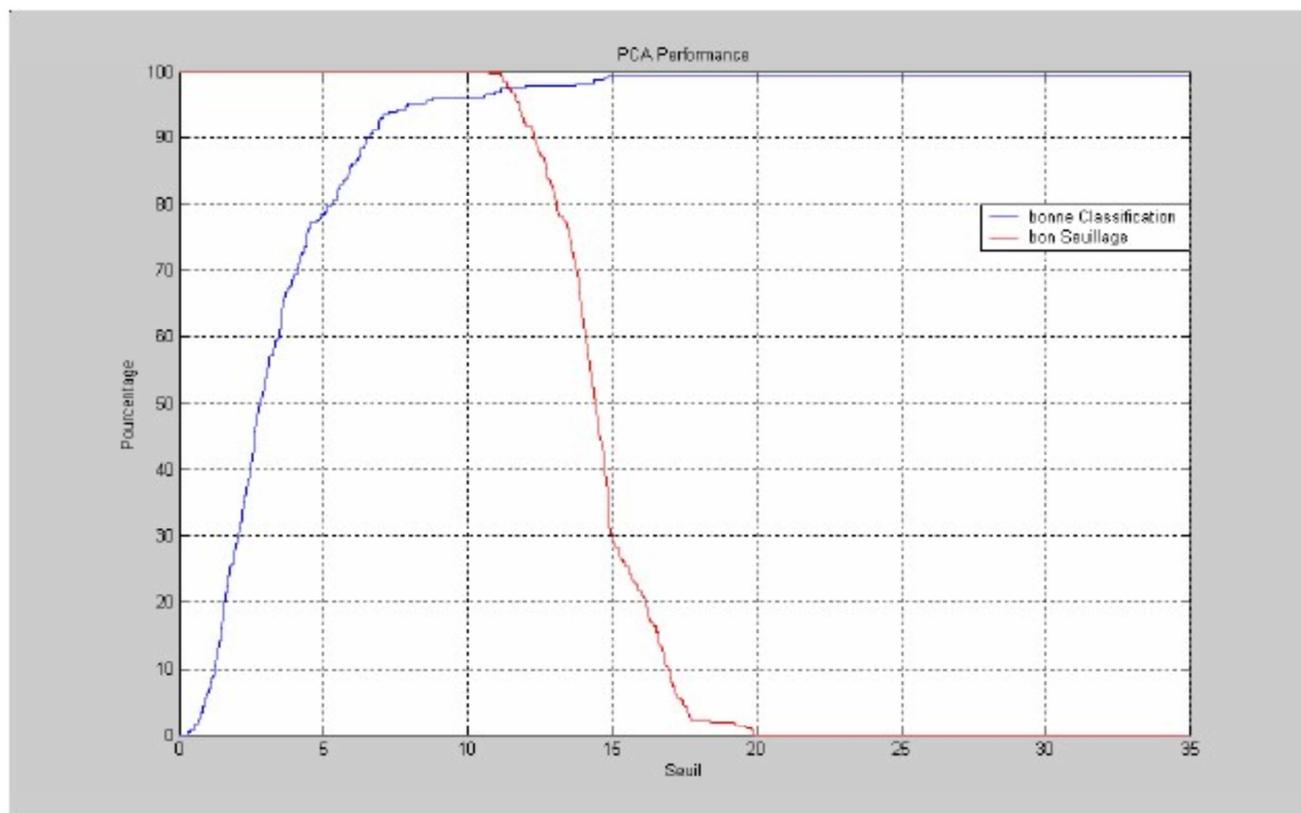


**FIG. III.20-** L'ensemble des images d'une personne de la base UMIST (on a pris 2 images sur 3 pour l'apprentissage et 1 image sur 3 comme image à identifier)

Et pour le seuillage des personnes inconnues, les performances sont :



**FIG. III.21-** Le graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil, de PCA sur UMIST



**FIG. III.22-** Le graphe des taux du bon seuillage et de la bonne classification en fonction de la valeur du seuil, de PCA sur UMIST

Pour une valeur de seuil = 11.12 on a : un pourcentage de bon seuillage des personnes inconnues égal à 100 % et un pourcentage de bonne classification des personnes connues égal à 96.5517 %.

Avec un petit seuil éliminant toutes les personnes inconnues, on arrive à un pourcentage de classification élevé, et ceci est dû à la faible différence de pose entre les images. Donc lorsqu'on calcule la différence entre les 2 vecteurs représentant l'image à identifier et l'image qui a été décidé par PCA comme étant l'image la plus proche de la base d'apprentissage, cette différence-là, est en fonction des identités des 2 images et pas en fonction de la pose (l'effet

de la pose est négligeable).

Distance = fct (différence d'identité)

Distance != fct (pose, éclairage)

Avec ces trois premières bases, on conclut que pour la pose, il faut avoir plusieurs angles de rotation plutôt que d'avoir, des poses frontales.

➤ Utilisation de la base HARVARD :

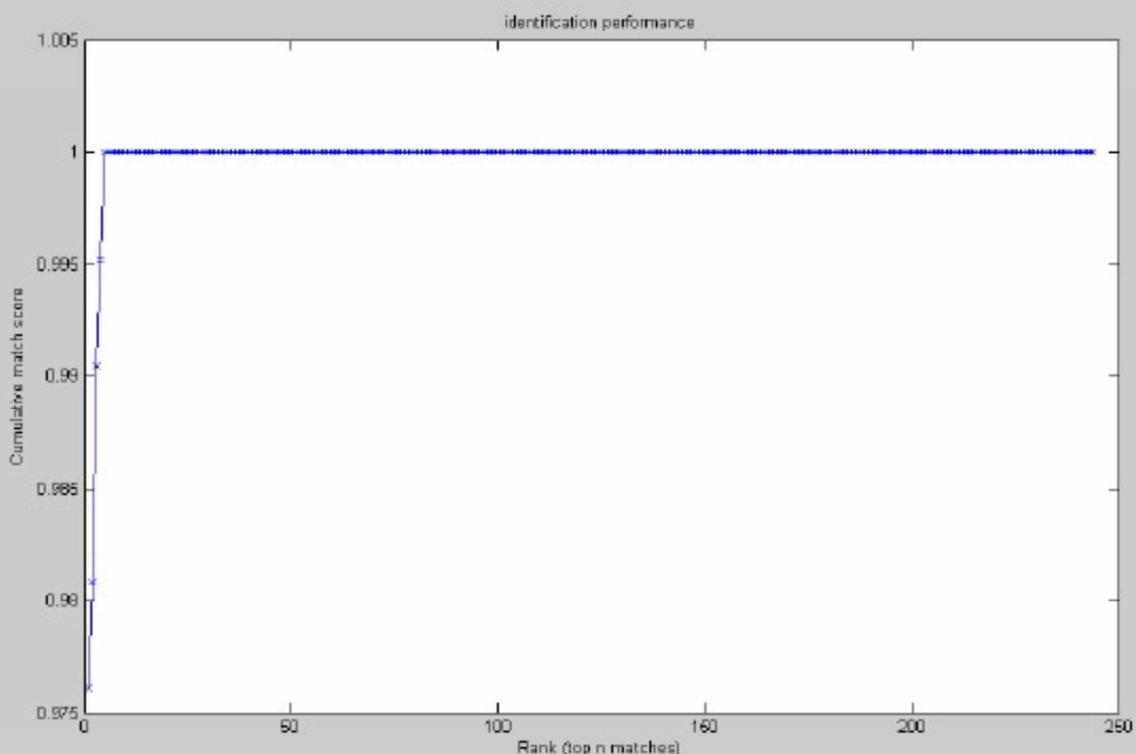
On a utilisé la base HARVARD pour vérifier l'effet néfaste de l'éclairage.

Au début, on a utilisé l'ensemble de la base (la totalité des images des 5 sets) pour avoir les performances globales de l'algorithme.

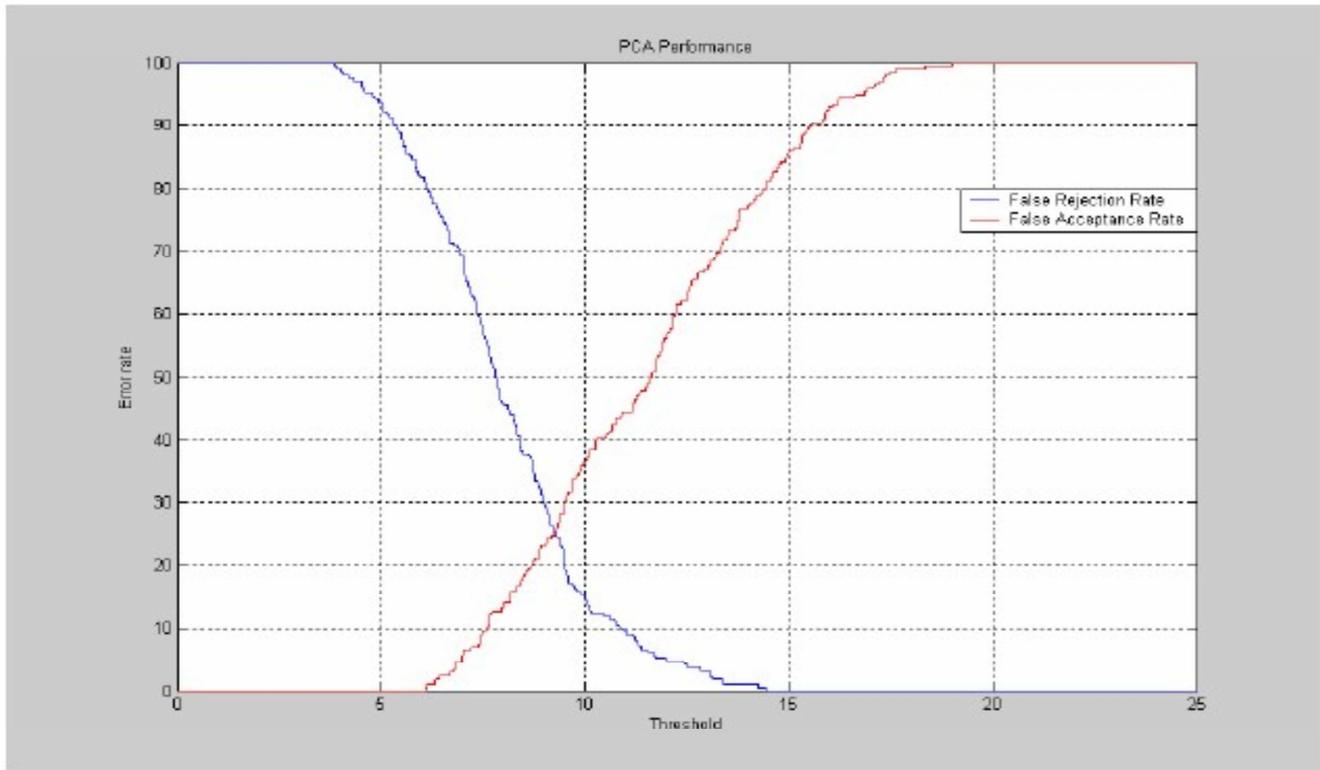
On a pris 244 images pour l'apprentissage (7 personnes), 209 images de test (7 personnes) et 198 images représentant les inconnus (3 personnes).

On a remarqué que l'étape de pré traitement améliore considérablement les performances, ce qui laisse à dire qu'elle est primordiale dans le cas où les conditions d'éclairage varient extrêmement.

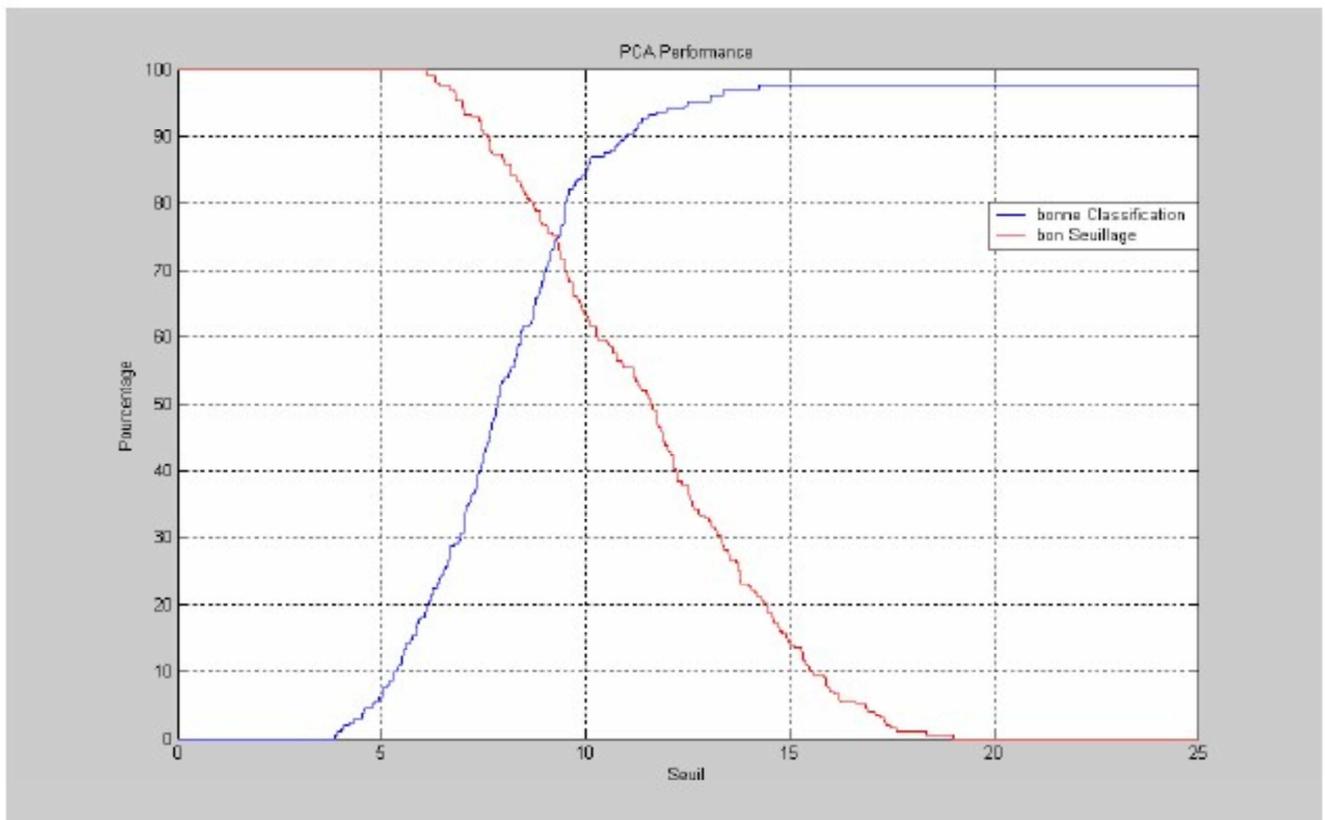
Les performances sur cette base sont :



**FIG. III.23- Le graphe des performances d'identification de PCA sur HARVARD**



**FIG. III.24- Le graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil, de PCA sur HARVARD**



**FIG. III.25- Le graphe des taux du bon seuillage et de la bonne classification en fonction de la valeur du seuil, de PCA sur HARVARD**

Dans la deuxième série de tests, PCA a été appliqué sur chaque Set séparément, et on a conclu qu'on atteint les meilleures performances dans le Set1 où les conditions sont les moins pressantes (les graphes des performances d'identification et du seuillage de chaque Set figurent dans l'annexe B).

- Set1: (FIG. B.1, FIG. B.2, FIG. B.3, annexe B)



**FIG. III.26- Les angles d'éclairage du Set1 de la base HARVARD :**  
*[o+15+00, o+15+15, o+15-15, o000+00, o000+15, o000-15]*

- Set2: (FIG. B.4, FIG. B.5, FIG. B.6, annexe B)



**FIG. III.27- Les angles d'éclairage du Set2 de la base HARVARD :**  
*[o+15+30, o+15-30, o+30+00, o+30+15, o+30+30, o+30-15, o+30-30, o000+30, o000-30]*

- Set3: (FIG. B.7, FIG. B.8, FIG. B.9, annexe B)



**FIG. III.28- Les angles d'éclairage du Set3 de la base HARVARD :**  
*[o+15+45, o+15-45, o+30+45, o+30-45, o+45+00, o+45+15, o+45+30, o+45+45, o+45-15, o+45-30, o+45-45, o000+45, o000-45]*

- Set4: (FIG. B.10, FIG. B.11, FIG. B.12, annexe B)



**FIG. III.29- Les angles d'éclairage du Set4 de la base HARVARD :**  
*[o+15+60, o+15-60, o+30+60, o+30-60, o+45+60, o+45-60, o+60+00, o+60+15,*  
*o+60+30, o+60+45, o+60+60, o+60-15, o+60-30, o+60-45, o+60-60, o000+60,*  
*o000-60]*

- Set5: (FIG. B.13, FIG. B.14, FIG. B.15, annexe B)



**FIG. III.30- Les angles d'éclairage du Set5 de la base HARVARD :**  
*[o+15+75, o+15-75, o+30+75, o+30-75, o+45+75, o+45-75, o+60+75, o+60-75,*  
*o+75+00, o+75+15, o+75+30, o+75+45, o+75+60, o+75+75, o+75-15, o+75-30, o+75-*  
*45, o+75-60, o+75-75, o000+75, o000-75]*

➤ Utilisation de la base FERET :

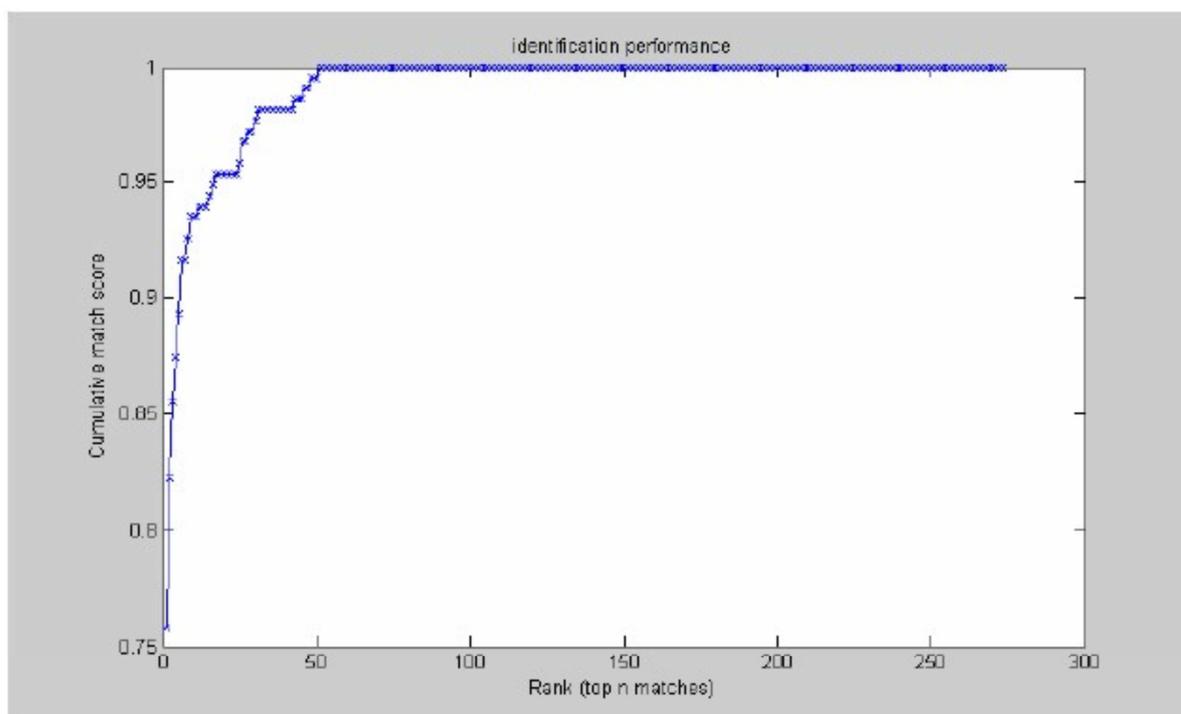
Les images utilisées résultent d'une conversion de l'espace couleur à l'espace noir et blanc, suivie d'un cropage et d'une normalisation de la taille.

L'étape de pré traitement améliore considérablement les performances vu la non uniformité des conditions d'éclairage.

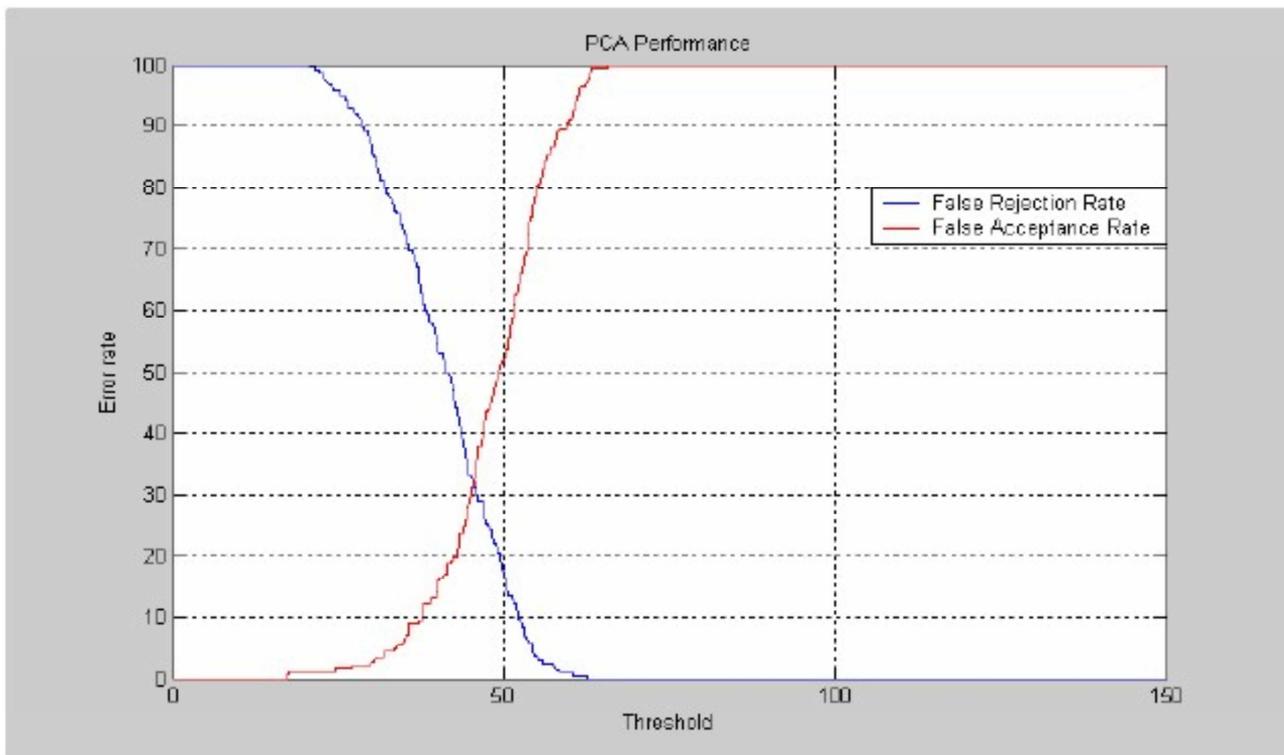
Cette base a été utilisée pour confirmer la baisse des performances, lors de conditions d'éclairage non identique.

On a fait les tests sur des images de pose : fa, fb, ql, qr, rb, et rc. La phase d'apprentissage contient 273 images, la phase de test 215 images, et 182 images représentent les inconnus.

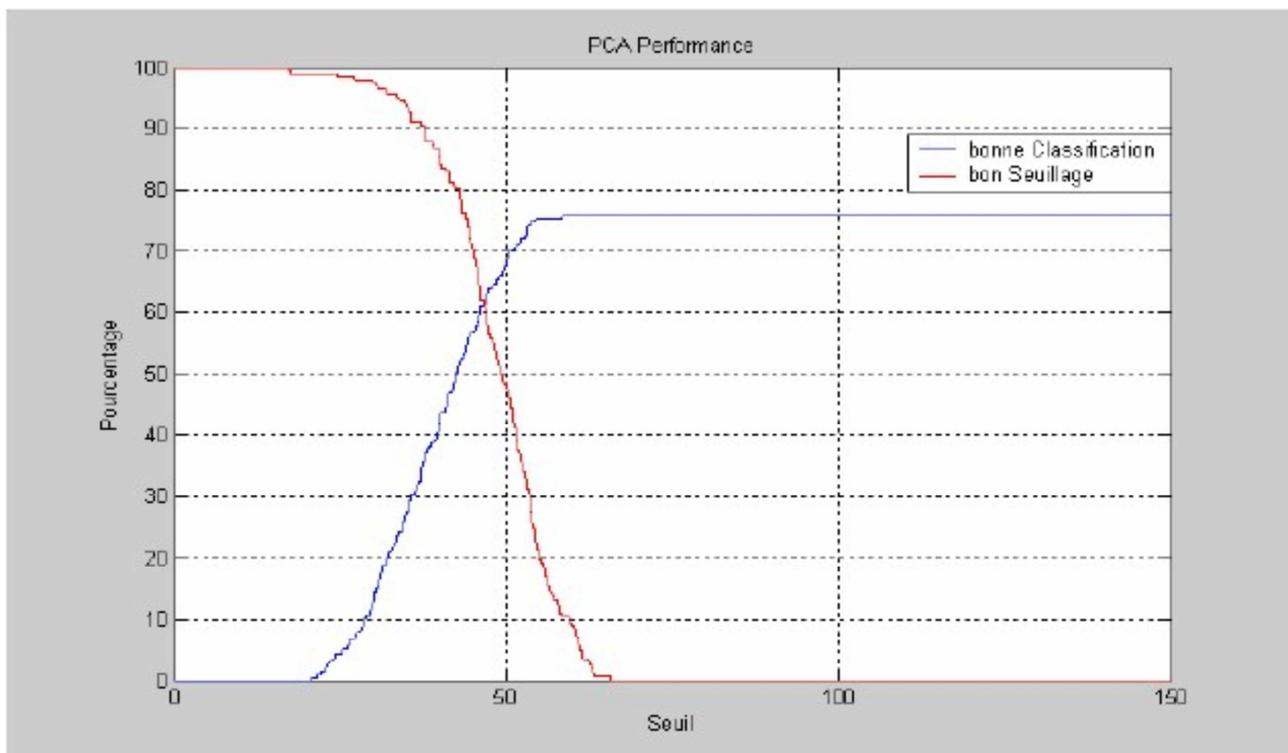
Les performances d'identification et de seuillage sont:



**FIG. III.31-** Le graphe des performances d'identification de PCA sur FERET



**FIG. III.32-** Le graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil, de PCA sur FERET



**FIG. III.33- Le graphe des taux du bon seuillage et de la bonne classification en fonction de la valeur du seuil, de PCA sur FERET**

➤ Utilisation de la base YALE :

La base YALE a été utilisée pour valider la dégradation des performances sous un éclairage non uniforme, et aussi pour vérifier ce qui est dit dans la littérature, sur l'effet néfaste des expressions faciales.

L'étape de pré traitement améliore les performances vu la non uniformité des conditions d'éclairage.

La phase d'apprentissage contient 70 images, la phase de test 40 images, et 55 images représentent les inconnus.

Les performances d'identification et de seuillage sont:

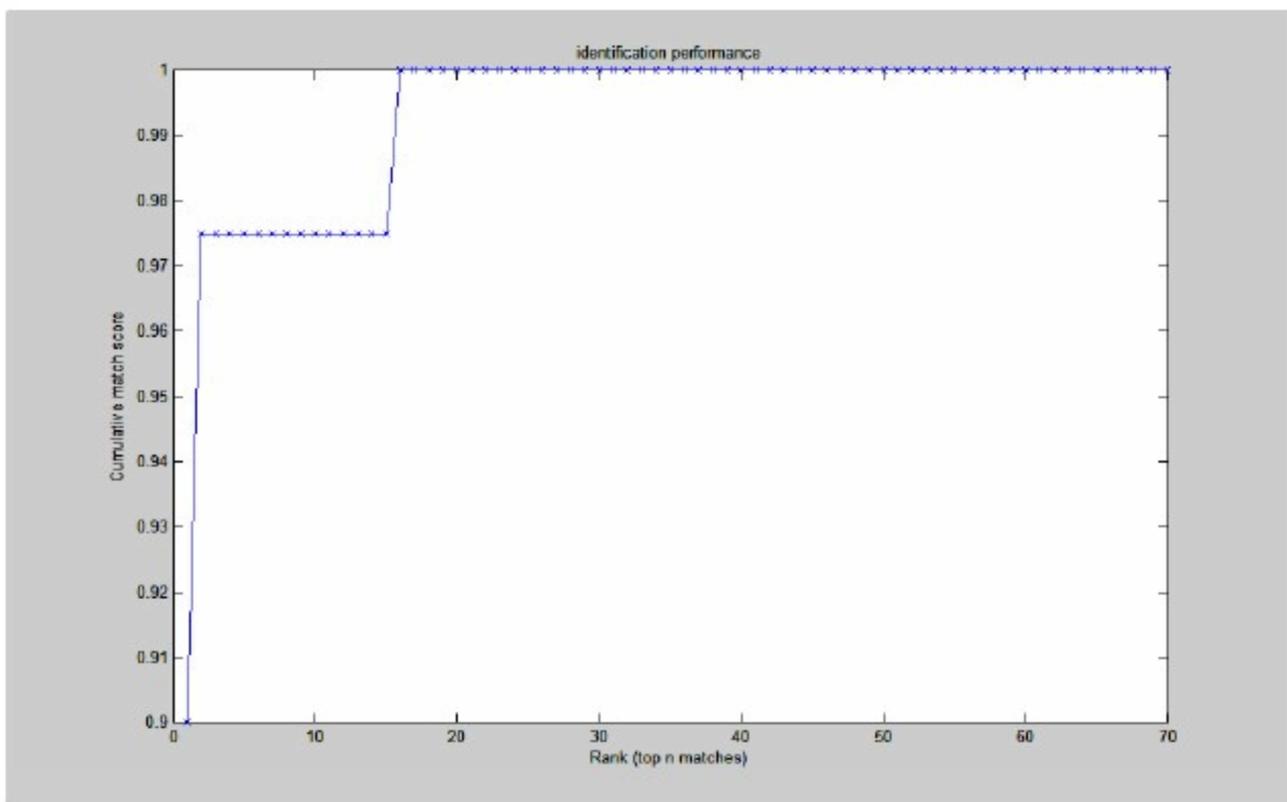


FIG. III.34- Le graphe des performances d'identification de PCA sur YALE

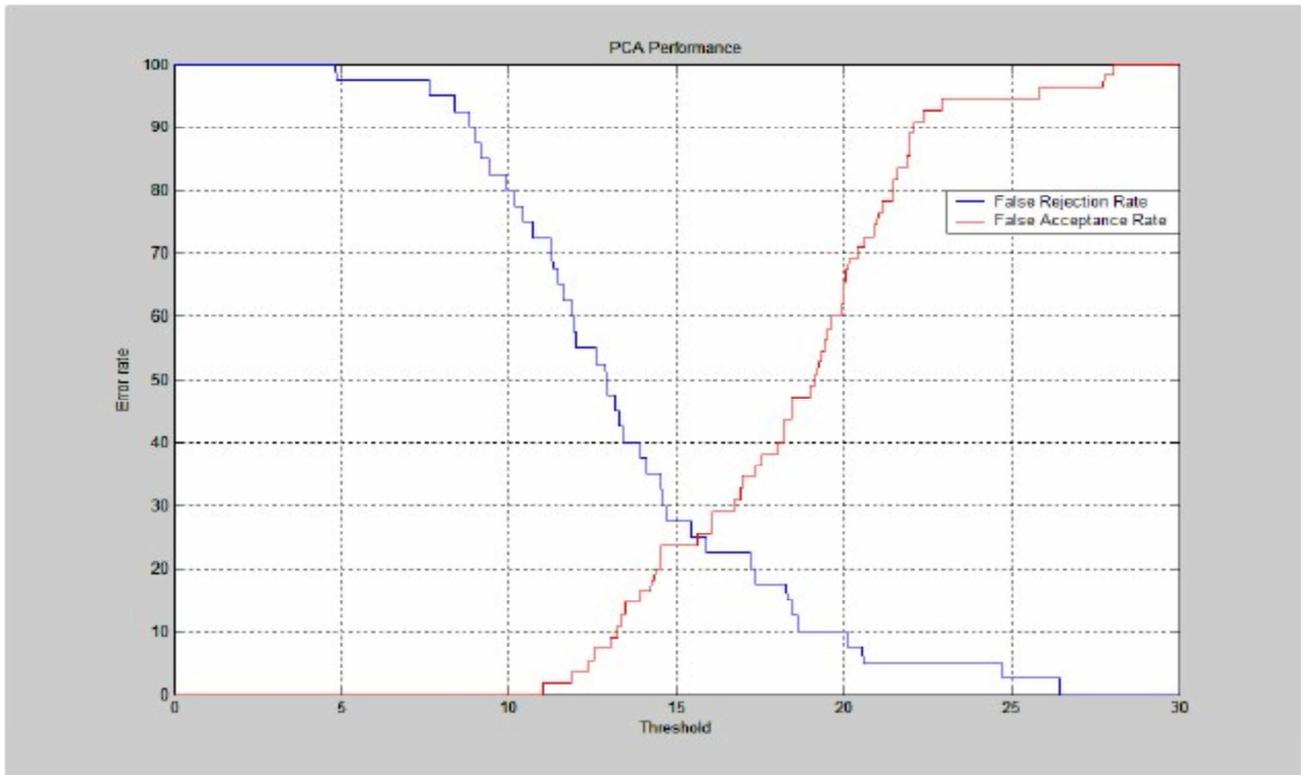


FIG. III.35- Le graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil, de PCA sur YALE

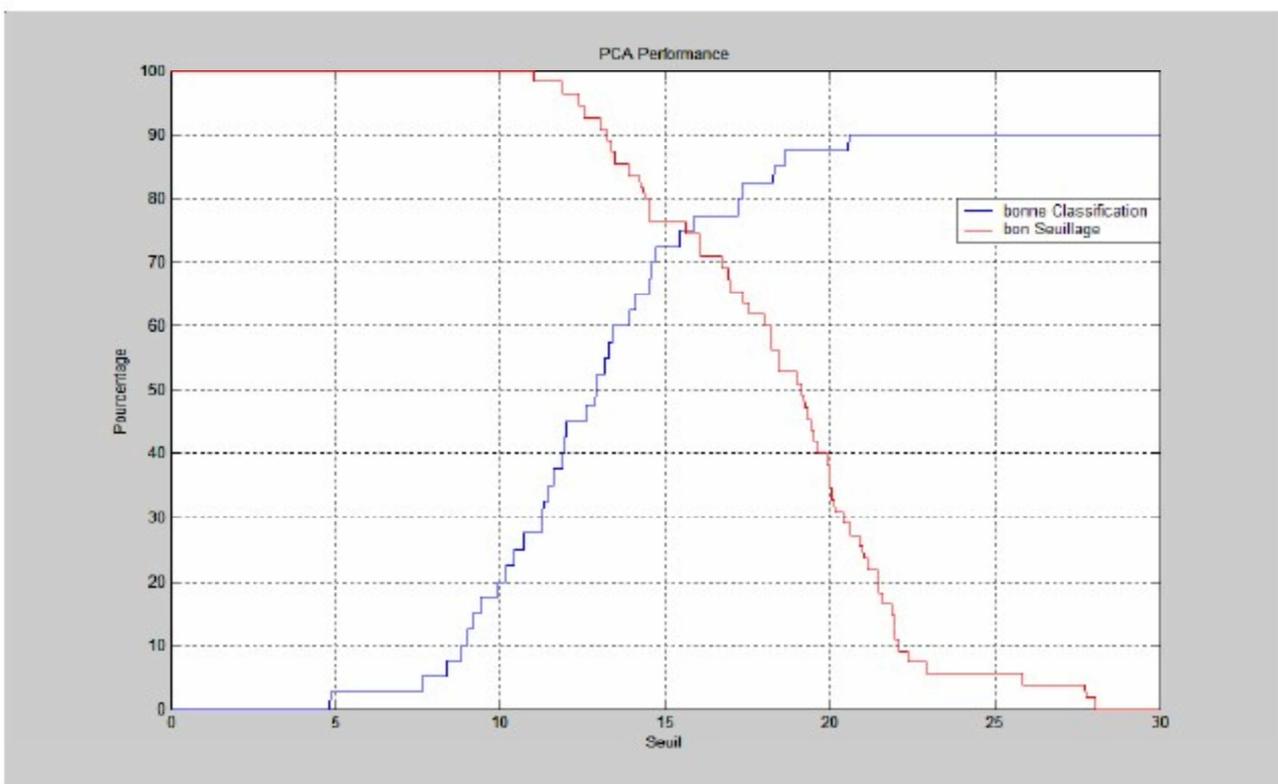


FIG. III.36- Le graphe des taux du bon seuillage et de la bonne classification en fonction de la valeur du seuil, de PCA sur YALE

### III.5.5 Tableau récapitulatif

Le tableau suivant regroupe l'ensemble des performances d'élimination des personnes inconnues.

	Base de données	Performances	Contraintes de la base
PCA	AT&T	Pour 100 % de bon seuillage on a <b>47.8261</b> % de bonne classification	<ul style="list-style-type: none"> <li>• L'application d'un cropage et d'une normalisation de la taille des images</li> <li>• L'utilisation d'images de face mal positionnées</li> <li>• La présence de lunettes et de composantes structurales</li> </ul>
		Pour 90 % de bon seuillage on a 69.5652 % de bonne classification	
		Pour 80 % de bon seuillage on a 80.0000 % de bonne classification	
	UMIST	Pour 100 % de bon seuillage on a <b>96.5517</b> % de bonne classification	<ul style="list-style-type: none"> <li>• L'éclairage uniforme</li> <li>• L'utilisation de plusieurs angles de pose pour chaque personne (très faible rotation entre les images)</li> </ul>
		90 % de bon seuillage on a 97.8448 % de bonne classification	
		Pour 80 % de bon seuillage on a 97.8448 % de bonne classification	
	HARVARD	Pour 100 % de bon seuillage on a <b>19.1388</b> % de bonne classification	<ul style="list-style-type: none"> <li>• La variation extrême des conditions d'éclairage</li> <li>• L'utilisation d'images de face Uniquement</li> </ul>
		Pour 90 % de bon seuillage on a 44.0191 % de bonne classification	
		Pour 80 % de bon seuillage on a 62.6795 % de bonne classification	
	Base d'images de face	Pour 100 % de bon seuillage on a <b>66.6667</b> % de bonne classification	<ul style="list-style-type: none"> <li>• La non uniformité des conditions d'éclairage</li> <li>• L'utilisation d'images de face Uniquement</li> </ul>
		Pour 90 % de bon seuillage on a 80.0000 % de bonne classification	
		Pour 80 % de bon seuillage on a 80.0000 % de bonne classification	
	FERET	Pour 100 % de bon seuillage on a <b>0</b> % de bonne classification	<ul style="list-style-type: none"> <li>• L'application d'un cropage et d'une normalisation de la taille</li> <li>• La présence de lunette et de composantes structurales</li> <li>• La prise des images en des</li> </ul>
		Pour 90 % de bon seuillage on a 36.2791 % de bonne classification	
		Pour 80 % de bon seuillage on a	

		51.3053 % de bonne classification	périodes espacées <ul style="list-style-type: none"> <li>• La non uniformité des conditions d'éclairage</li> <li>• L'utilisation des poses : fa, fb, ql, qr, rb, et rc</li> </ul>
	YALE	Pour 100 % de bon seuillage on a <b>27.5000</b> % de bonne classification	<ul style="list-style-type: none"> <li>• La présence de lunettes et de composantes structurelles</li> <li>• Plusieurs expressions faciales</li> <li>• Plusieurs conditions d'éclairage</li> <li>• L'utilisation d'images de face uniquement</li> </ul>
		Pour 90 % de bon seuillage on a 55.0000 % de bonne classification	
		Pour 80 % de bon seuillage on a 65.0000 % de bonne classification	

**TAB. III.10- L'ensemble des performances d'élimination des personnes inconnues, de PCA sur AT&T, UMIST, HARVARD, FERET, YALE et une base d'images de face**

### III.5.6 Détermination du nombre de poses à utiliser dans l'apprentissage

On constate que les performances sont relatives à la base, puisque les conditions dans lesquelles on la construit influent directement sur les performances (les graphes des performances d'identification et de seuillage différent d'une base à une autre).

Et on remarque que l'idéal est d'avoir une base qui respecte les conditions suivantes : un éclairage uniforme pendant les deux phases d'apprentissage et de reconnaissance, et la prise de différentes poses pour chaque personne pendant la phase d'apprentissage, avec à chaque fois une légère rotation.

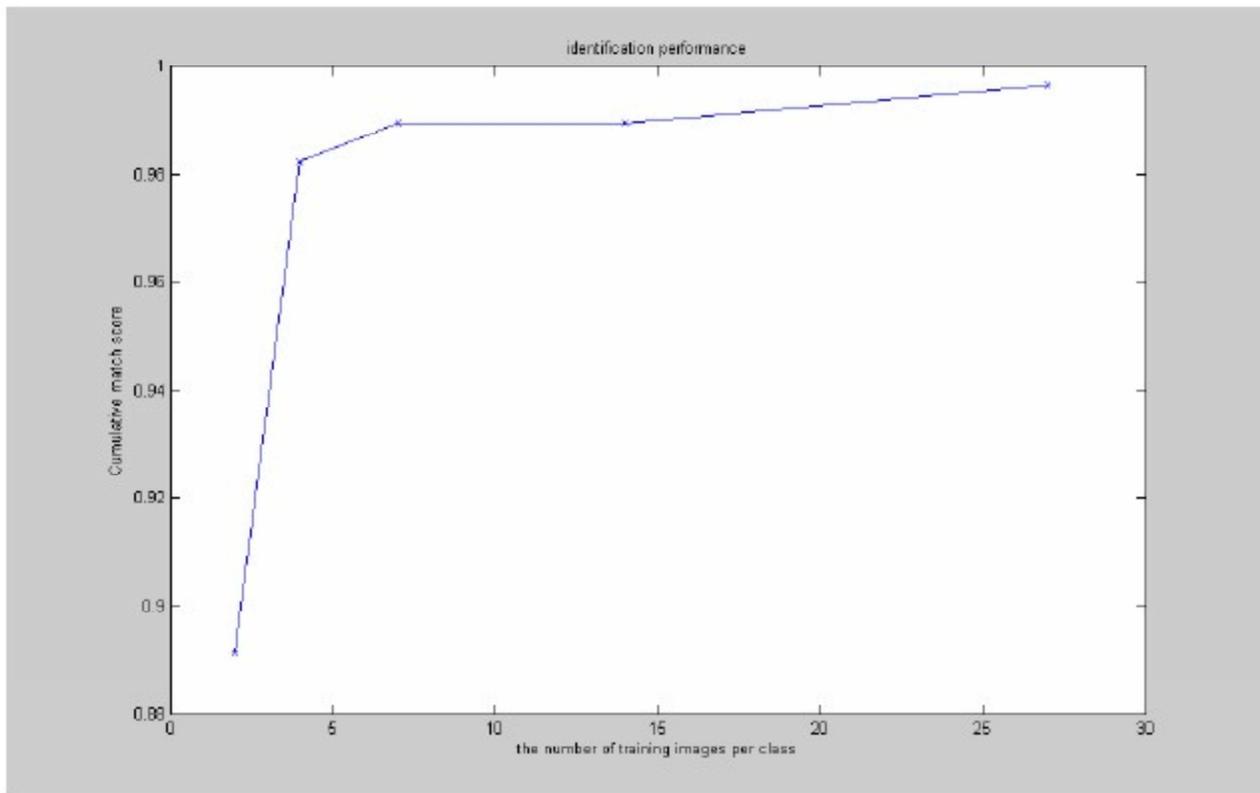
Si le nombre de poses par personne est très grand, alors on a les meilleures performances, mais en contrepartie ça augmente la mémoire de stockage des informations significatives des images d'apprentissage, et aussi le temps de reconnaissance. C'est pour cela qu'on a essayé de déterminer un nombre de poses adéquat, qui soit petit et qui offre en même temps des performances acceptables.

Durant des tests, on a utilisé 11 personnes (classes) de la base UMIST. On a pris 286 images de test (26 images par personne), et on a fait varier le nombre d'images d'apprentissage par personne.

La variation du nombre d'images d'apprentissage représente en fait, la variation de rotation

entre les poses de chaque personne, puisque lorsqu'on augmente le nombre d'images, en réalité on ajoute de nouvelles poses situées entre les poses antérieures, donc c'est comme si on diminuait la rotation entre les prises d'images d'apprentissage.

Pour chaque nombre d'images d'apprentissage, on a calculé le taux de bonne identification, et l'ensemble des résultats est représenté par le graphe suivant :



**FIG. III.37- Le taux de bonne identification en fonction du nombre d'images d'apprentissage par personne**

L'axe horizontal de ce graphe représente le nombre d'images d'apprentissage par personne (on a pris 2, 4, 7, 14, 27), et l'axe vertical quant à lui, représente le taux de bonne identification. Ce graphe montre bien que tant que le nombre de poses par personne augmente, tant que le taux de bonne identification augmente.

Les tests ont été fait avec 10 images d'apprentissage par personne (FIG. C.1, FIG. C.2, FIG. C.3, annexe C), et avec 15 images (FIG. C.4, FIG. C.5, FIG. C.6, annexe C), et on a trouvé qu'avec 15 images, on a des performances acceptables, et c'est le nombre qu'on a choisi (les graphes des performances d'identification et du seuillage avec 10 images d'apprentissage par

personne, et avec 15 images, figurent dans l'annexe C).

Le tableau suivant regroupe différents taux de bonne classification des sujets connus, en fonction du taux de la bonne élimination des personnes inconnues, en prenant à chaque fois 10 et 15 images d'apprentissage par personne :

	10 images d'apprentissage	15 images d'apprentissage
100% de bon seuillage	89.1667% de bonne classification	<b>90.8257%</b> de bonne classification
90% de bon seuillage	92.5000% de bonne classification	93.2110% de bonne classification
80% de bon seuillage	93.5000% de bonne classification	93.9450% de bonne classification

***TAB. III.11- Les performances d'identification et de seuillage, pour 10 et 15 images d'apprentissage par personne***

### **III.6 Conclusion**

L'ensemble des tests effectués a permis de conclure, qu'avec l'utilisation de PCA sous un éclairage uniforme tout en ayant plusieurs poses par personne, permet d'obtenir des performances optimales.

## Conclusion générale

Vu la naissante nécessité d'utiliser des applications de contrôle d'accès, la reconnaissance de visages a émergé comme un secteur actif de recherches, enjambant des disciplines telles que le traitement d'image, l'identification de modèle, et la vision d'ordinateur.

Dû à sa nature facile à utiliser, la reconnaissance de visage restera un outil puissant malgré l'existence d'autres méthodes biométriques de reconnaissance.

Durant ces dernières années, de nombreuses méthodes ont été proposées, dont plusieurs ont été appliquées, avec succès. Le choix d'une méthode devrait être basé sur les conditions spécifiques de chaque application.

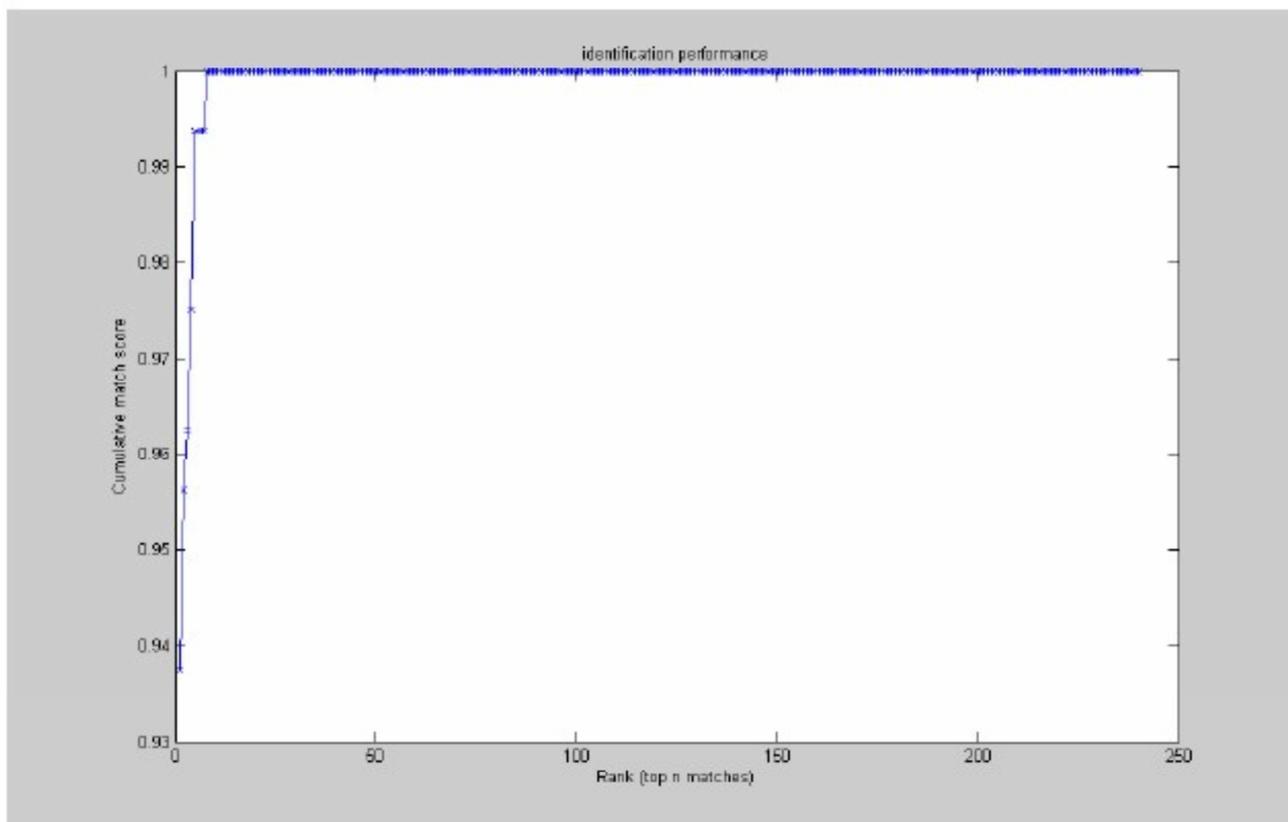
Parmi toutes ces méthodes, l'algorithme de PCA reste une des approches les plus fiables et les plus simples, offrant à la fois de bonnes performances de classification des personnes connues et de bonnes performances d'élimination des personnes inconnues.

Une étape cruciale dans l'élaboration de tous systèmes, et l'évaluation. Une étape qui permet de dégager les paramètres et les conditions idéales à suivre.

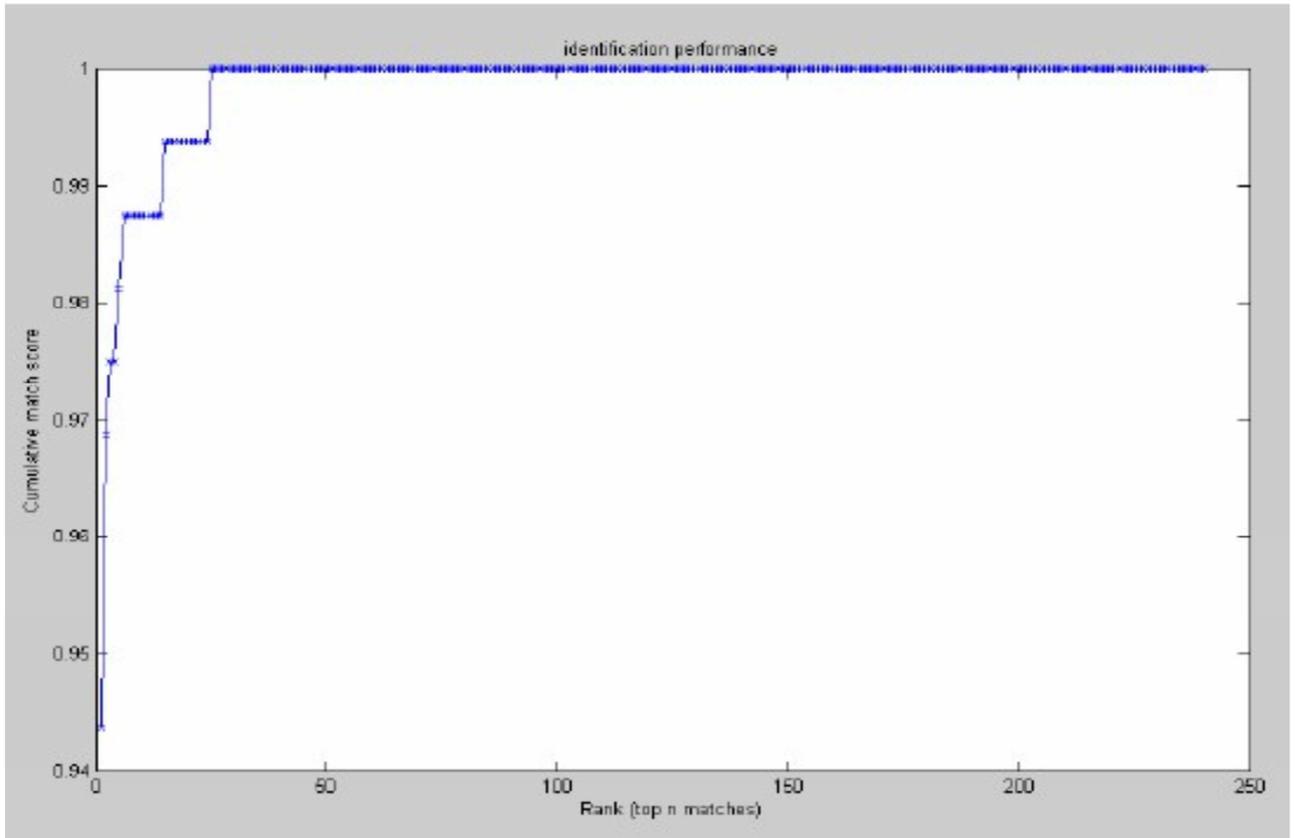
Malgré tous les progrès qui ont été réalisés, les problèmes de pose et d'éclairage, et de l'identification dans des environnements extérieurs restent des challenges qui susciteront les efforts des chercheurs.

Actuellement, il y'a une nouvelle tendance qui arrive et qui commence à susciter les efforts, c'est le multimodale, dans lequel on combine plusieurs technologies biométriques, ou plusieurs algorithmes de reconnaissance, ou on utilise divers classificateurs pondérés pour essayer d'améliorer les performances de reconnaissance. Parmi les perspectives ouvertes à ce projet, l'utilisation d'autres méthodes de reconnaissance comme HMM et EBG, et la combinaison avec d'autre technologie biométriques comme l'empreinte digitale ou l'iris.

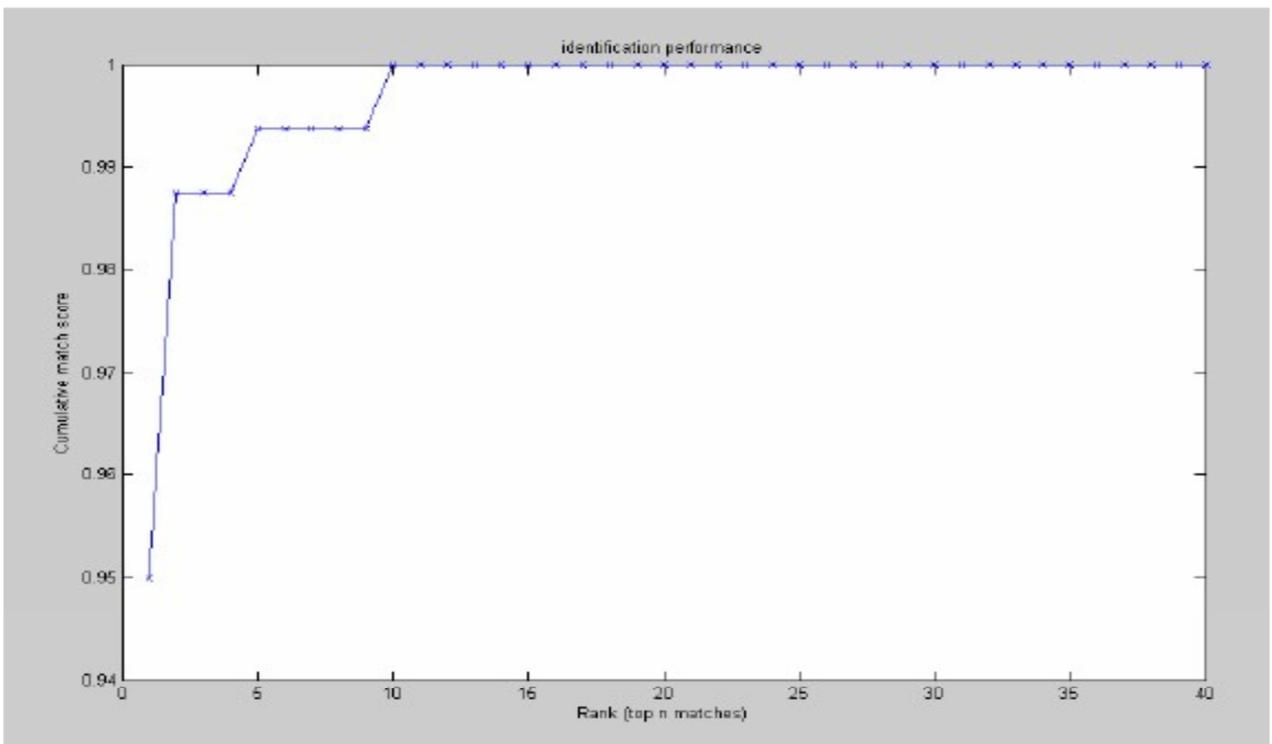
Annexe A Les graphes des performances d'identification des algorithmes PCA, DF LDA, SVM, Template-matching, la distance de Hausdorff, les CROSS RATIO, PCA appliquée aux régions des yeux et PCA appliquée aux régions des yeux et de la bouche



*FIG. A.1- Le graphe des performances d'identification de PCA*

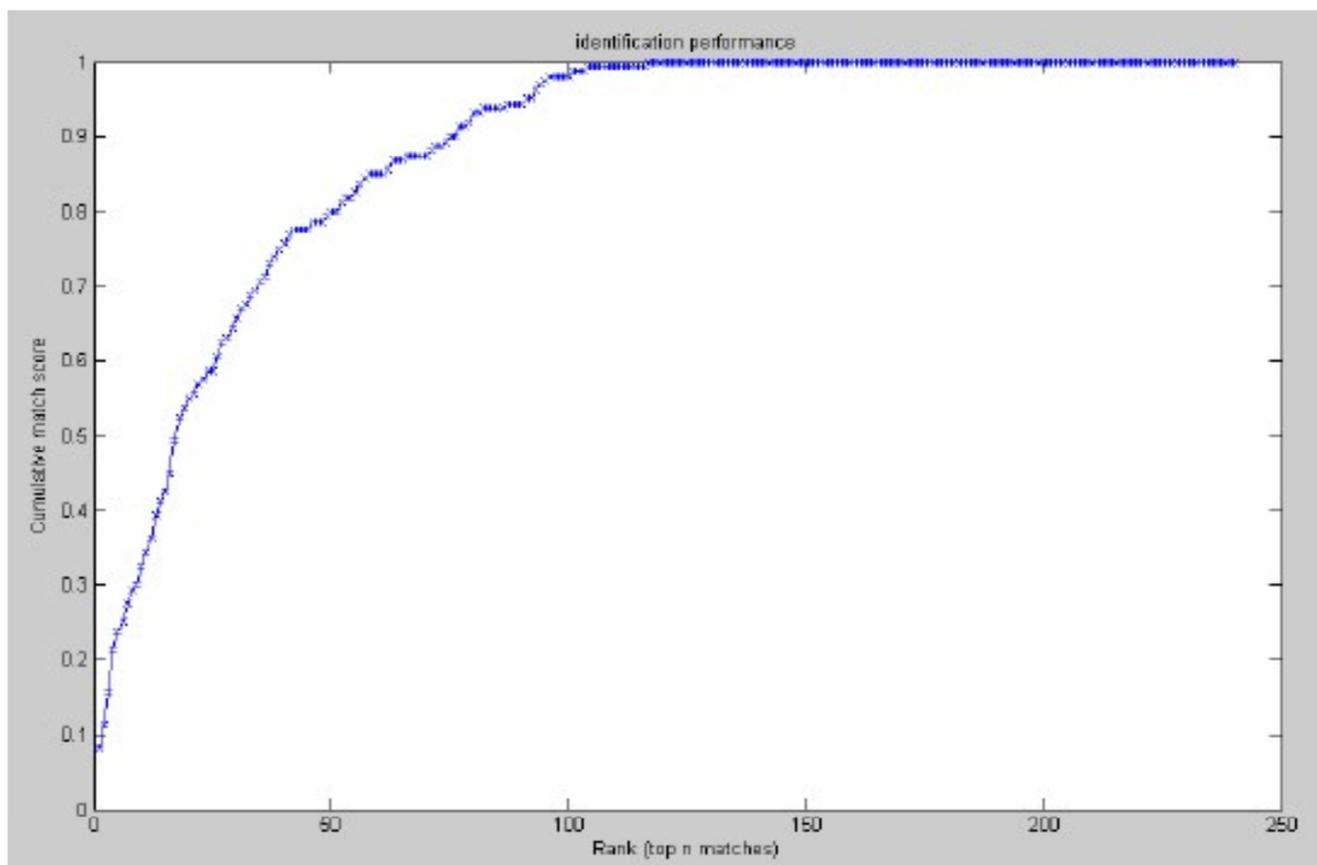


*FIG. A.2- Le graphe des performances d'identification de DF-LDA*

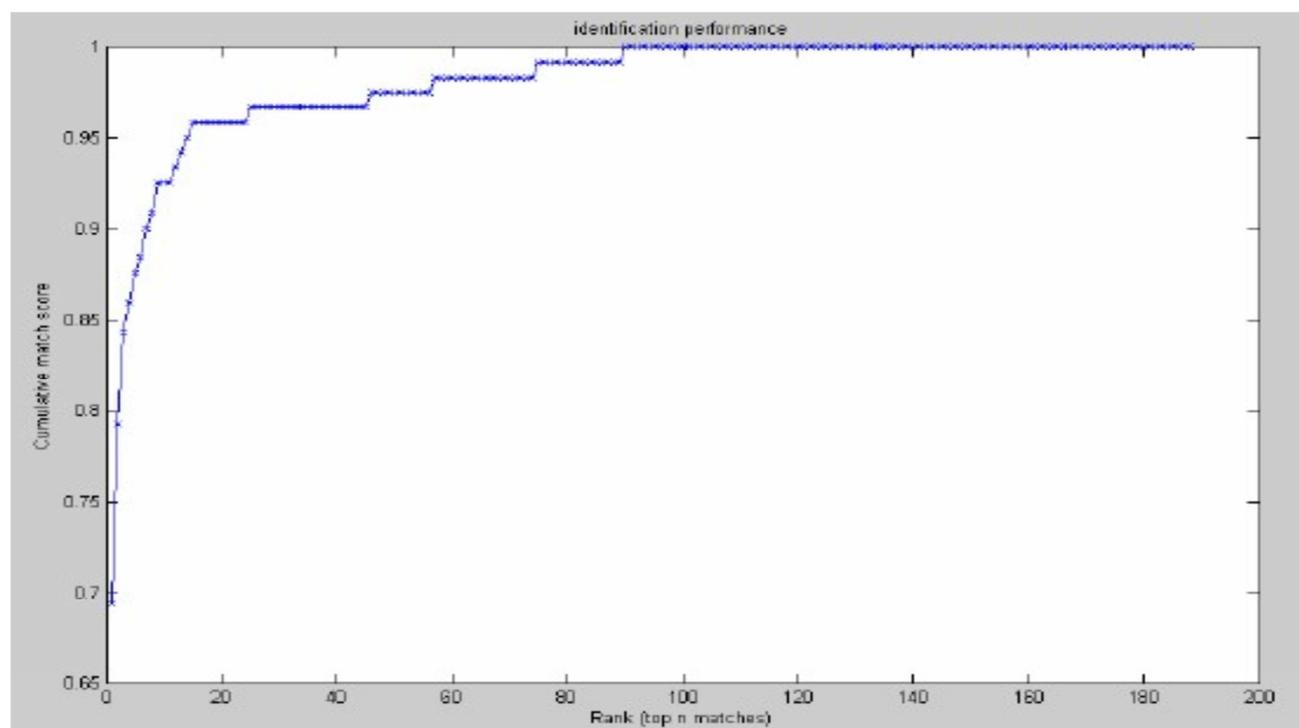


*FIG. A.3- Le graphe des performances d'identification de SVM*

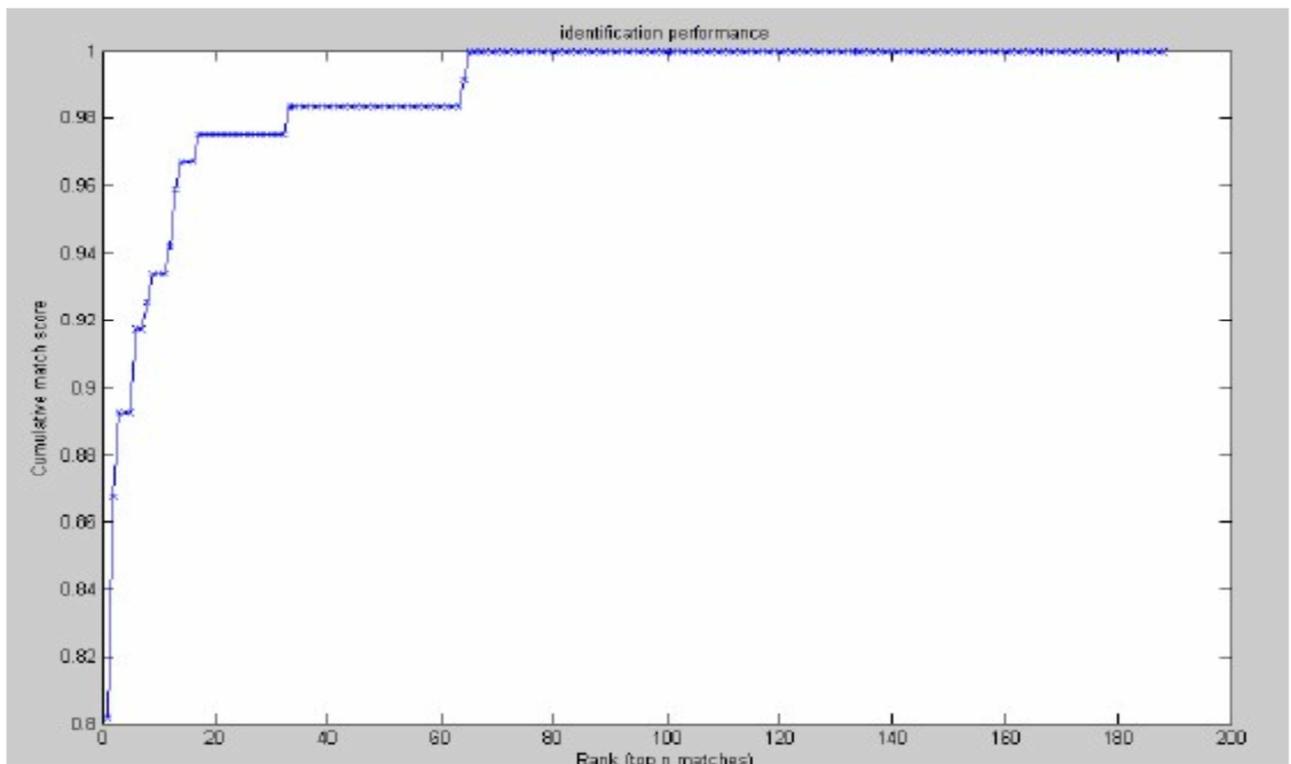




*A.6- Le graphe des performances d'identification des CROSS RATIO*

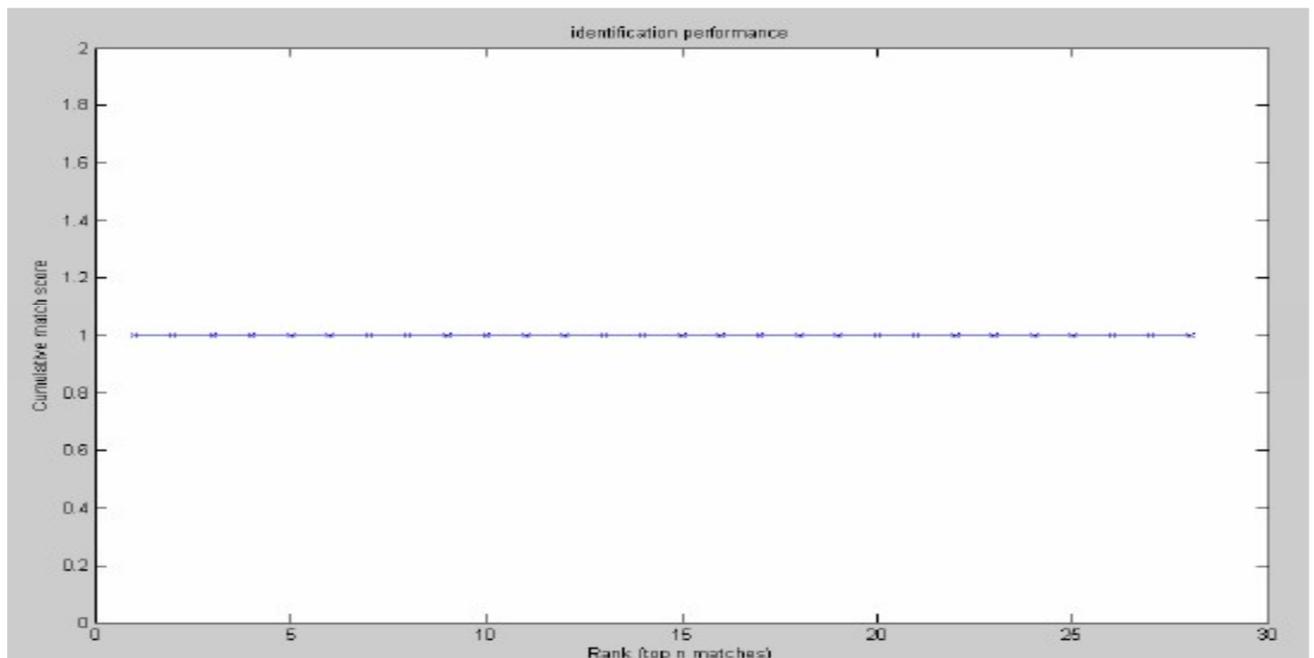


*FIG. A.7- Le graphe des performances d'identification de PCA appliquée aux régions des yeux*

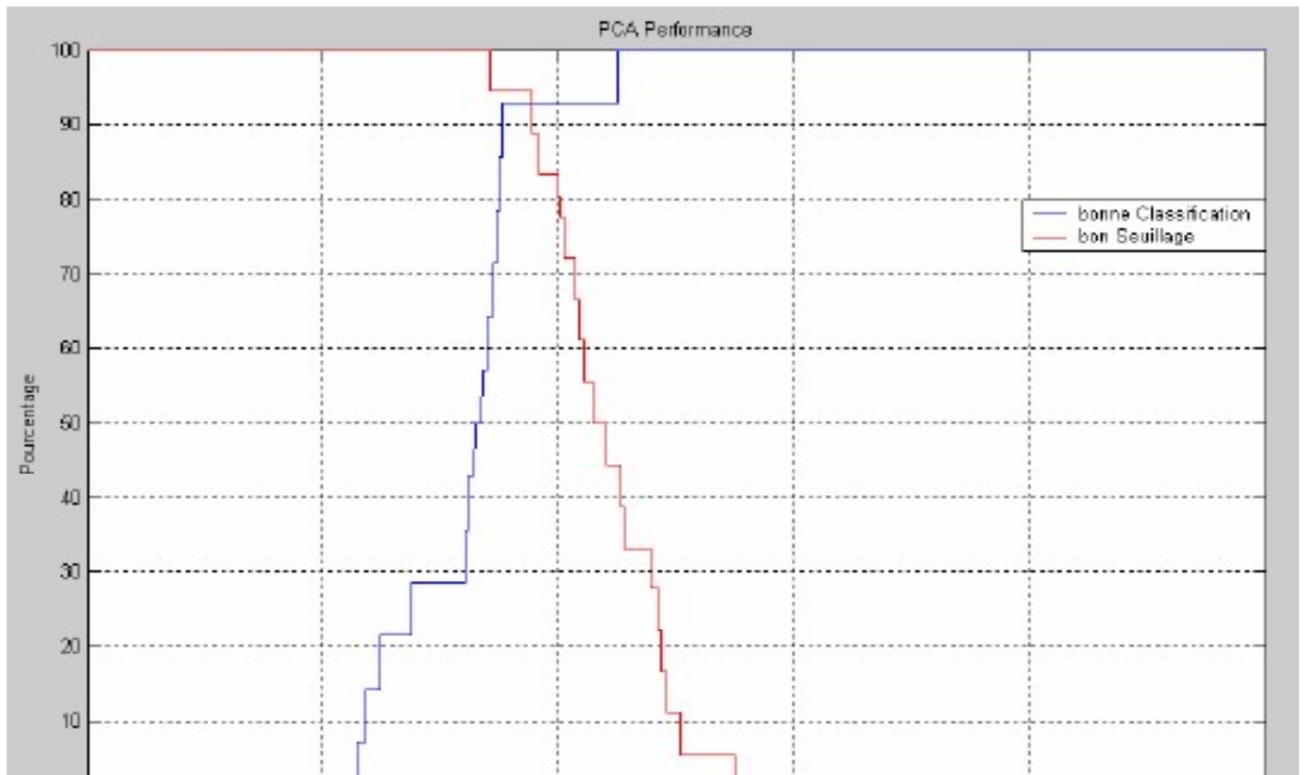
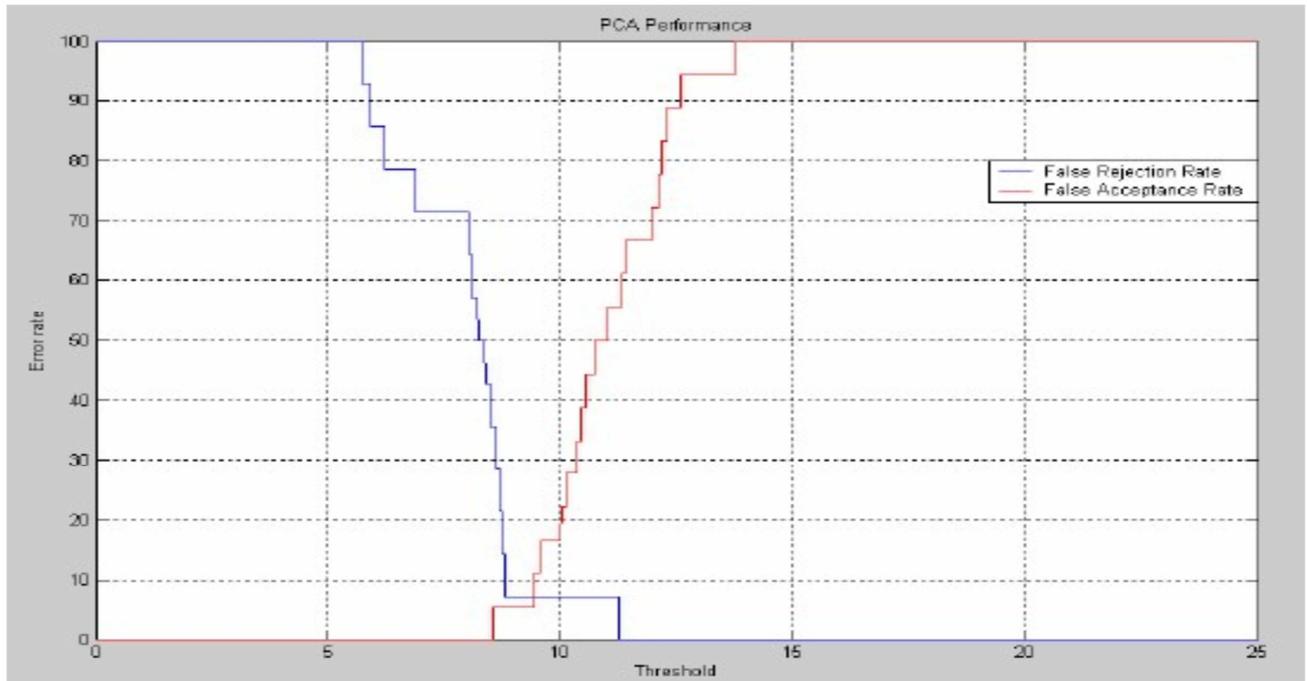


**FIG. A.8-** *Le graphe des performances d'identification de PCA appliquée aux régions des yeux et de la bouche*

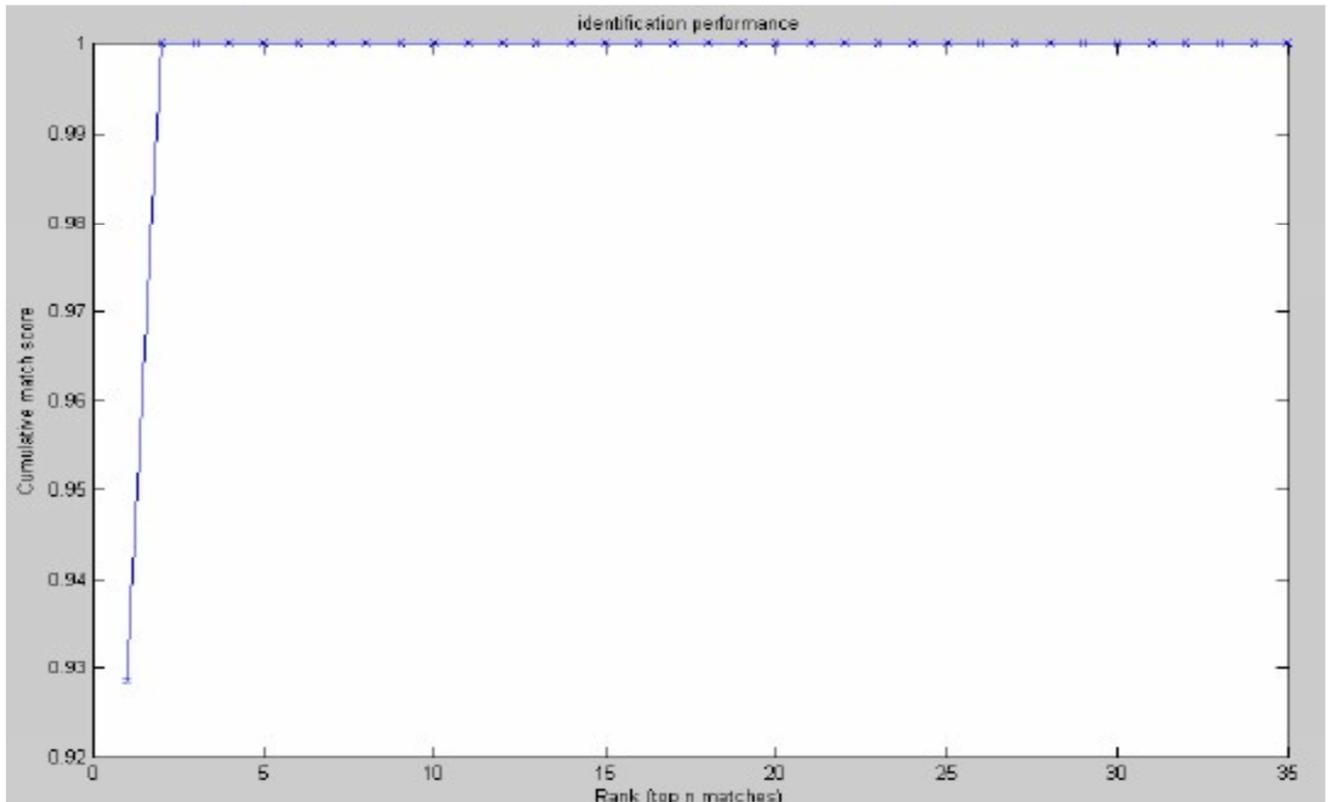
Annexe B Les graphes des performances d'identification et du seuillage de PCA sur chaque Set de la base Harvard



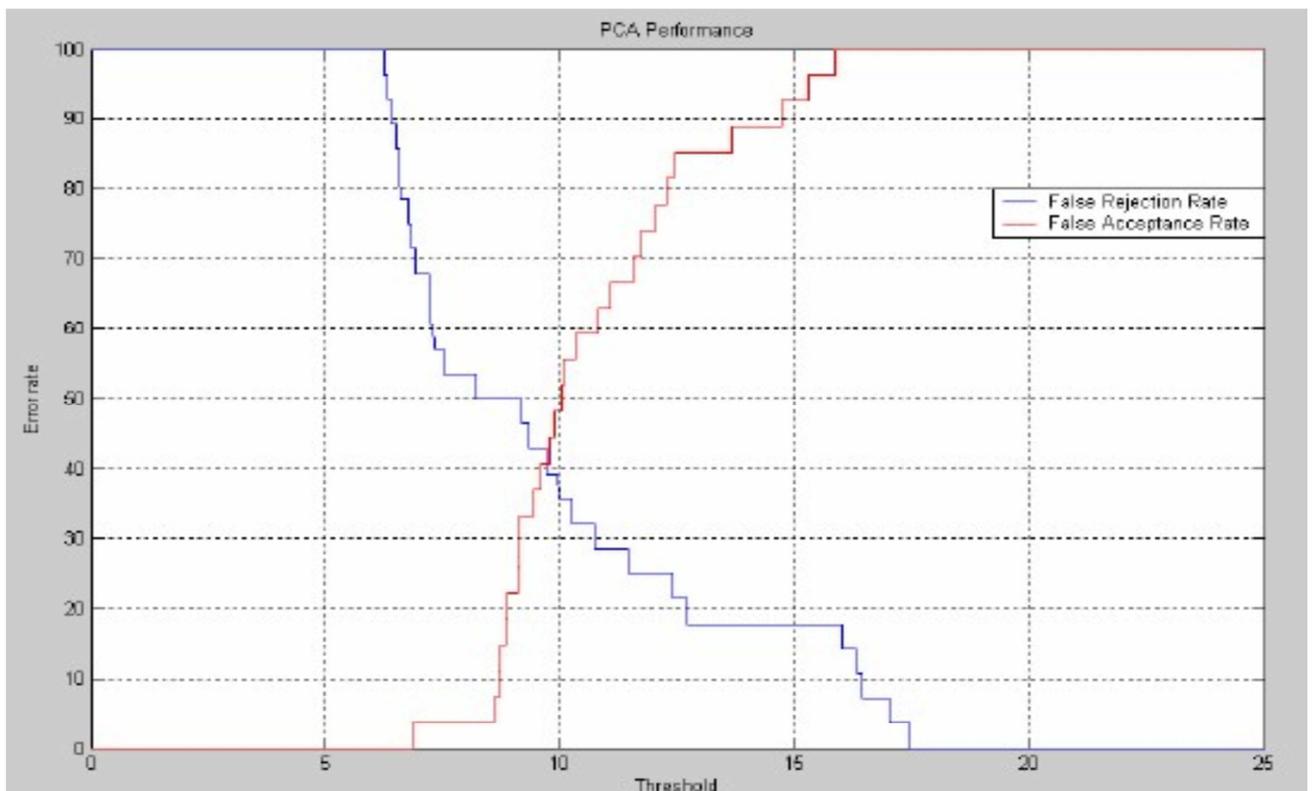
**FIG. B.1-** *Le graphe des performances d'identification de PCA sur le Set1 de la base HARVARD*



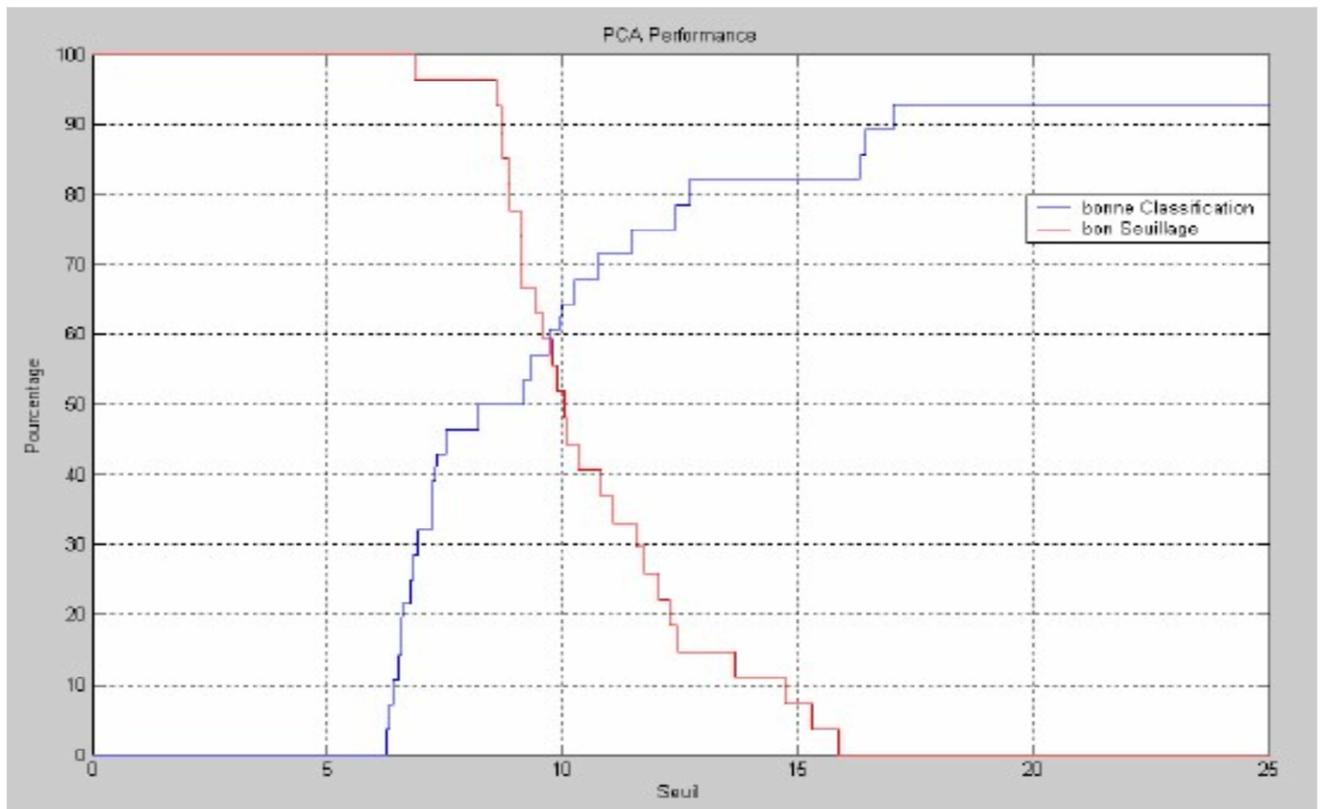
**FIG. B.3-** Le graphe des taux du bon seuillage et de la bonne classification en fonction de la valeur du seuil, de PCA sur le Set1 de la base HARVARD



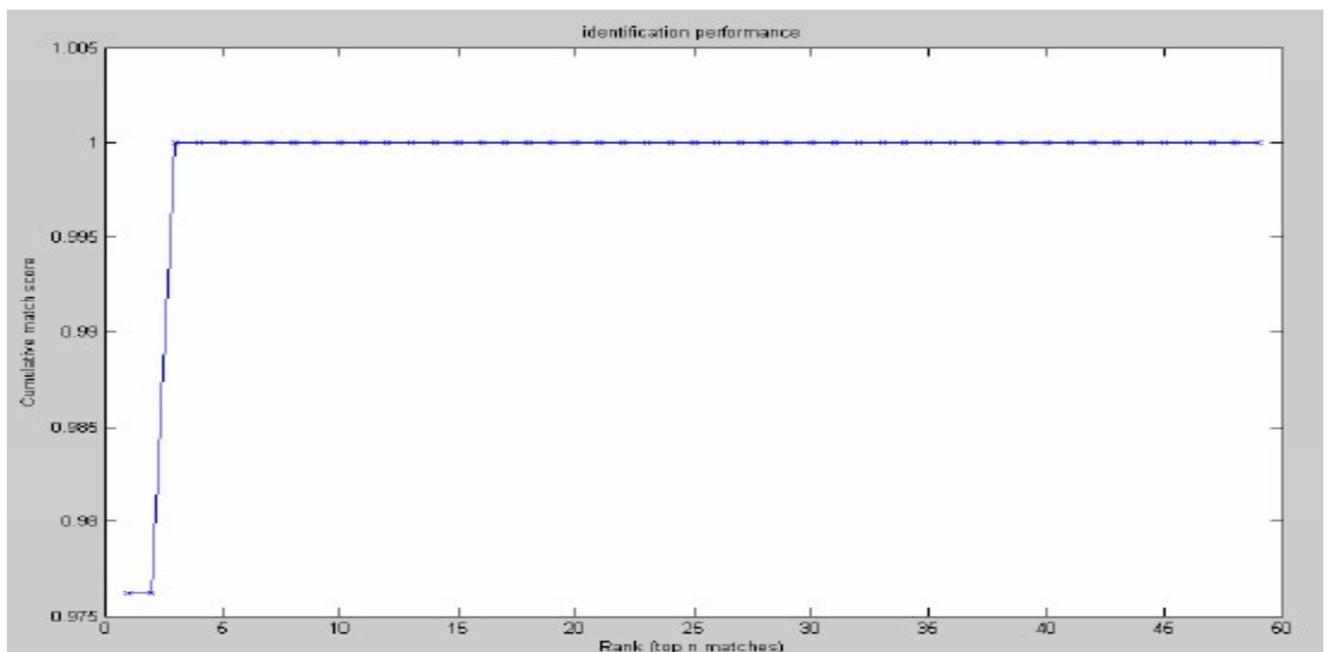
**FIG. B.4-** Le graphe des performances d'identification de PCA sur le Set2 de la base HARVARD



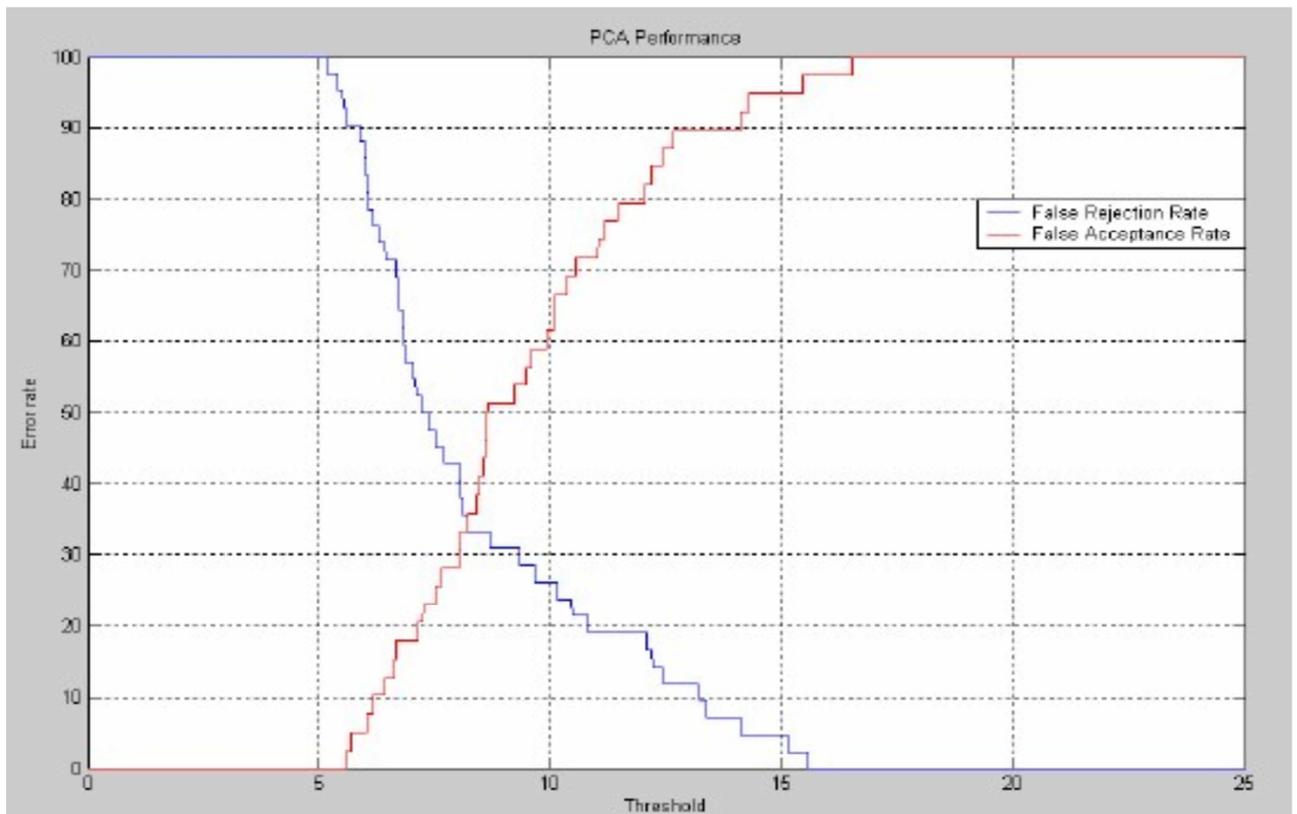
**FIG. B.5-** Le graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil, de PCA sur le Set2 de la base HARVARD



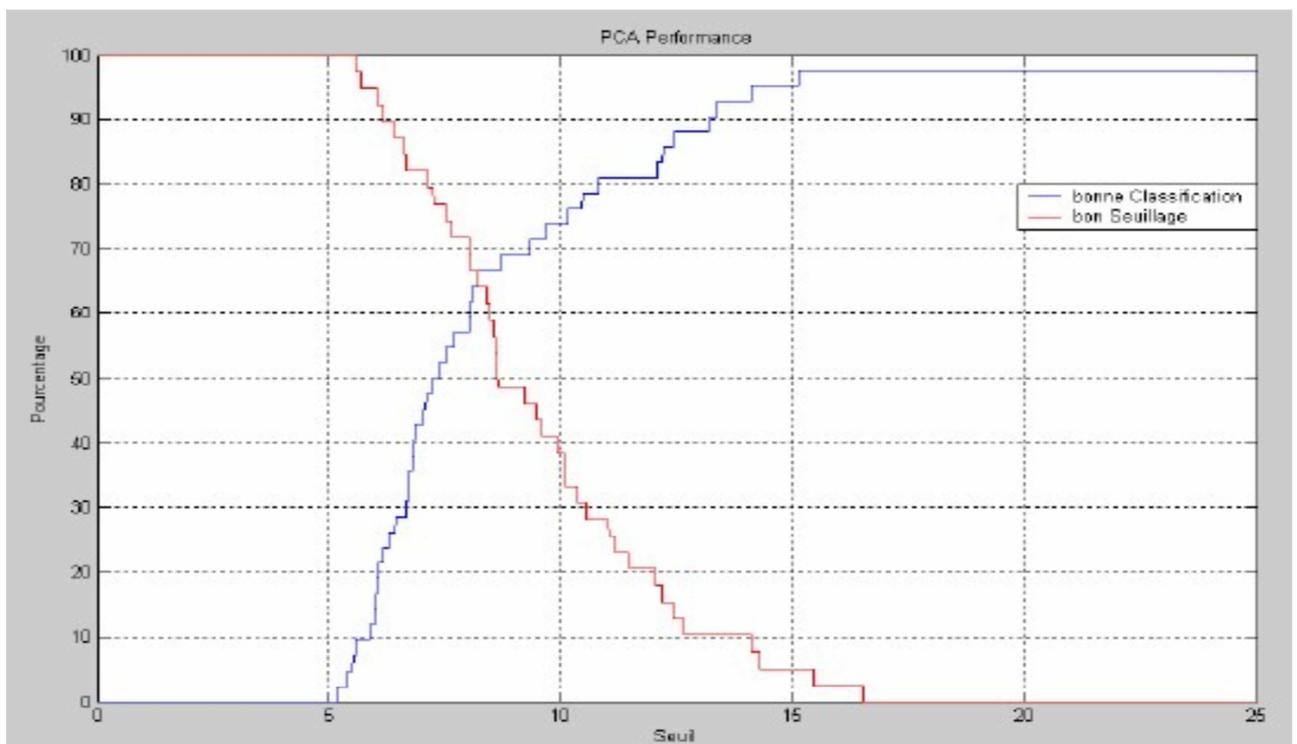
**FIG. B.6-** Le graphe des taux du bon seuillage et de la bonne classification en fonction de la valeur du seuil, de PCA sur le Set2 de la base HARVARD



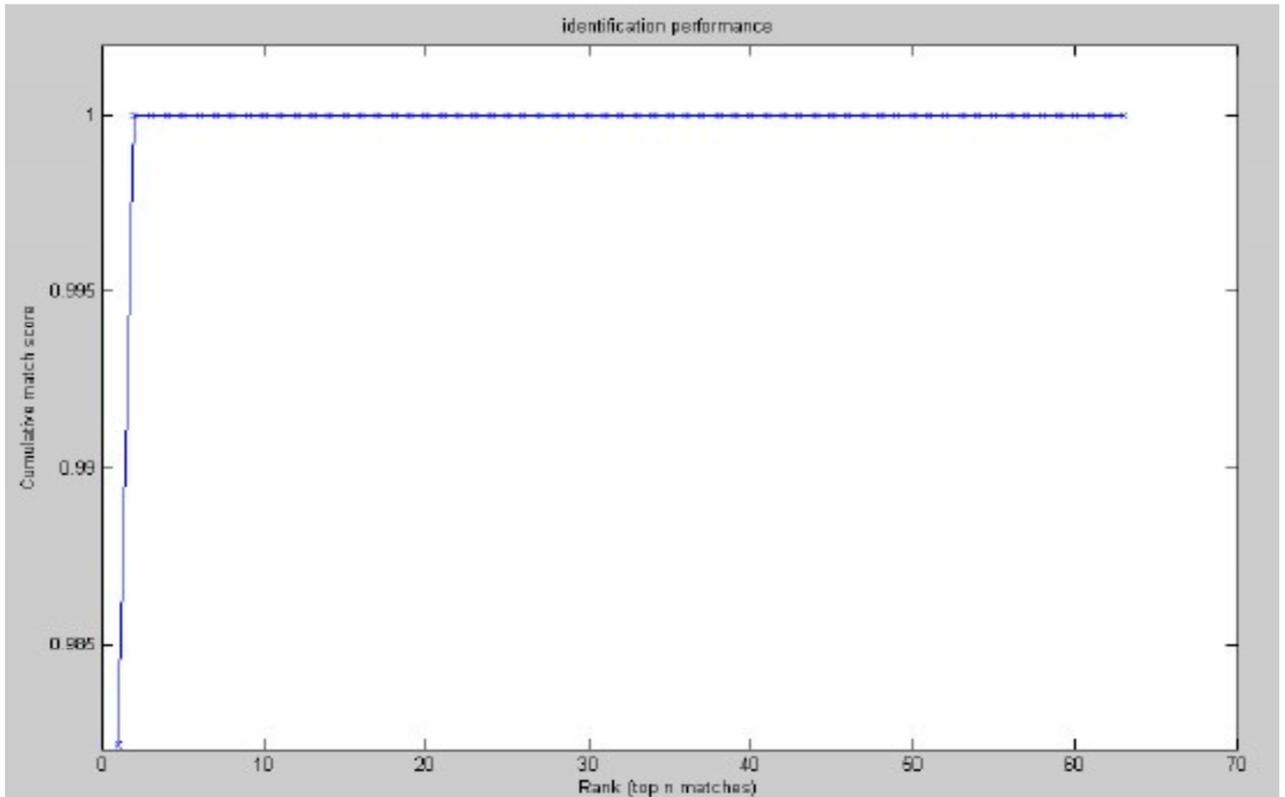
**FIG. B.7-** Le graphe des performances d'identification de PCA sur le Set3 de la base HARVARD



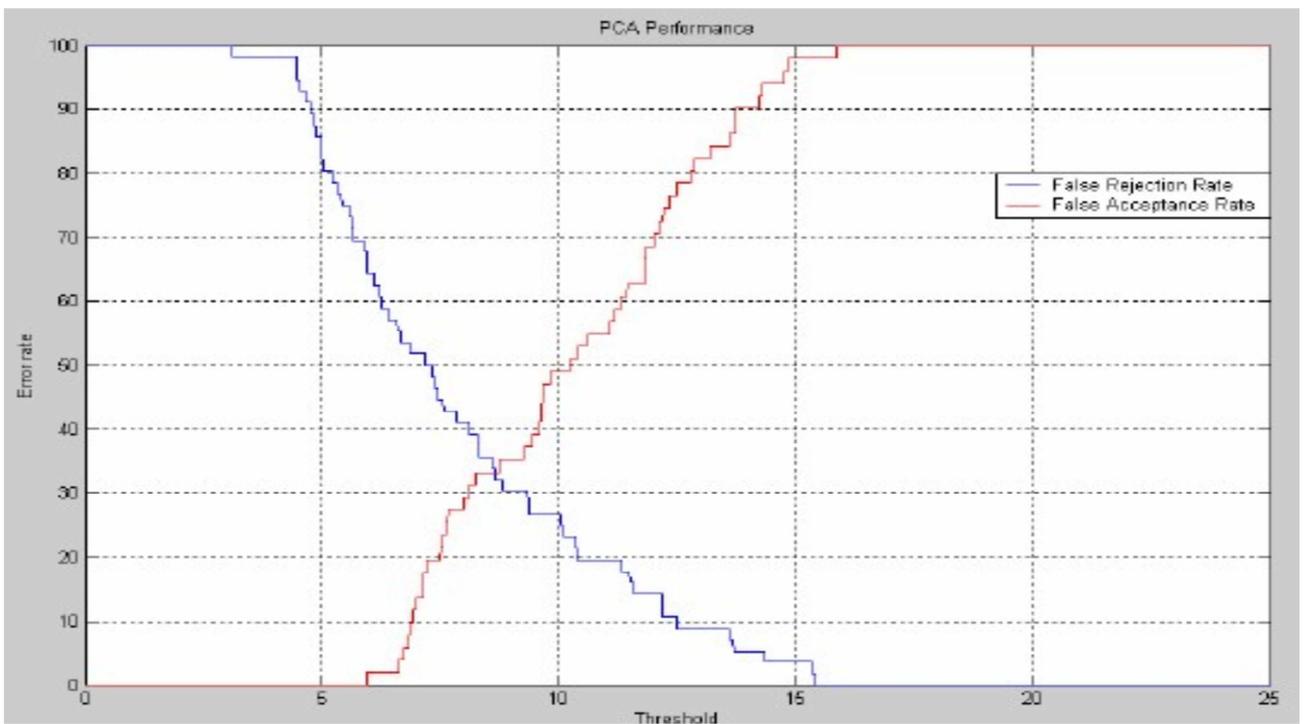
**FIG. B.8-** Le graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil, de PCA sur le Set3 de la base HARVARD



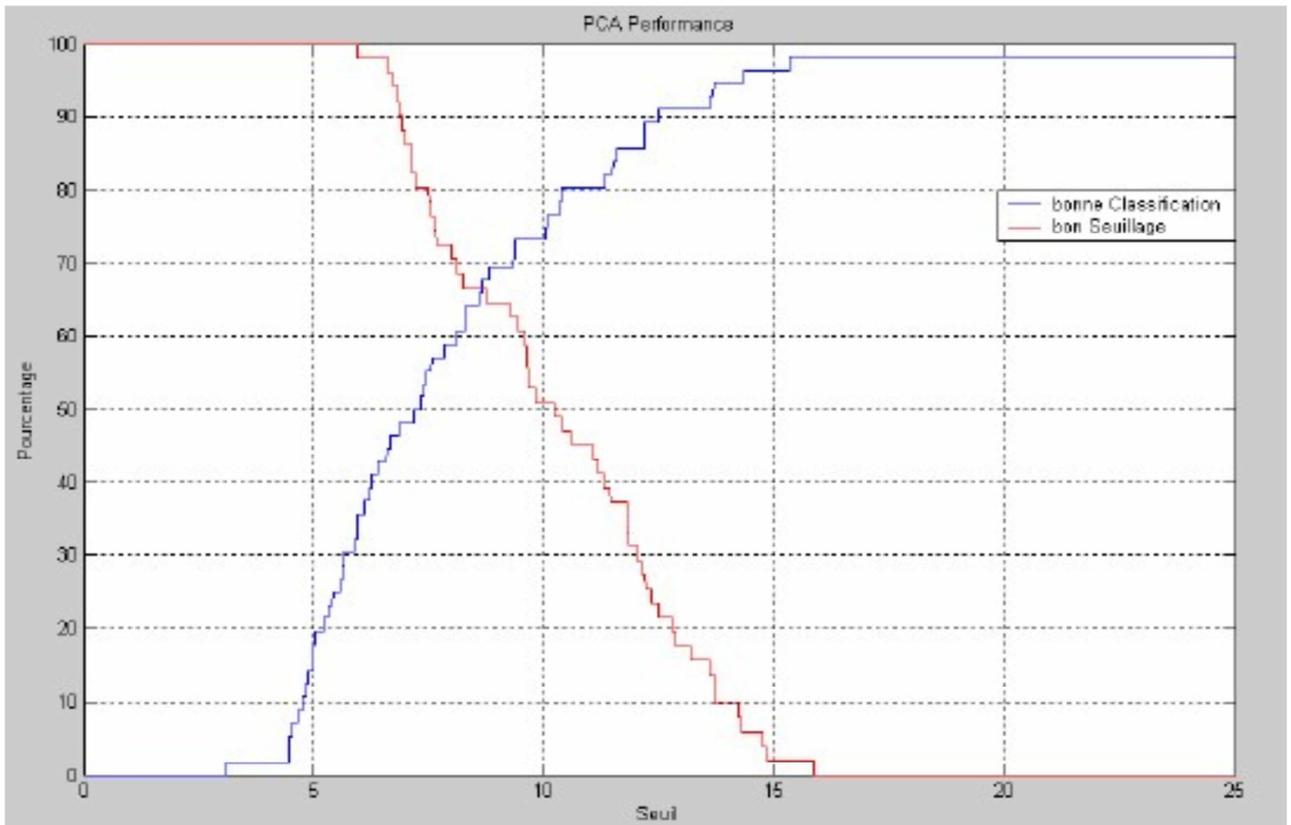
**FIG. B.9-** Le graphe des taux du bon seuillage et de la bonne classification en fonction de la valeur du seuil, de PCA sur le Set3 de la base HARVARD



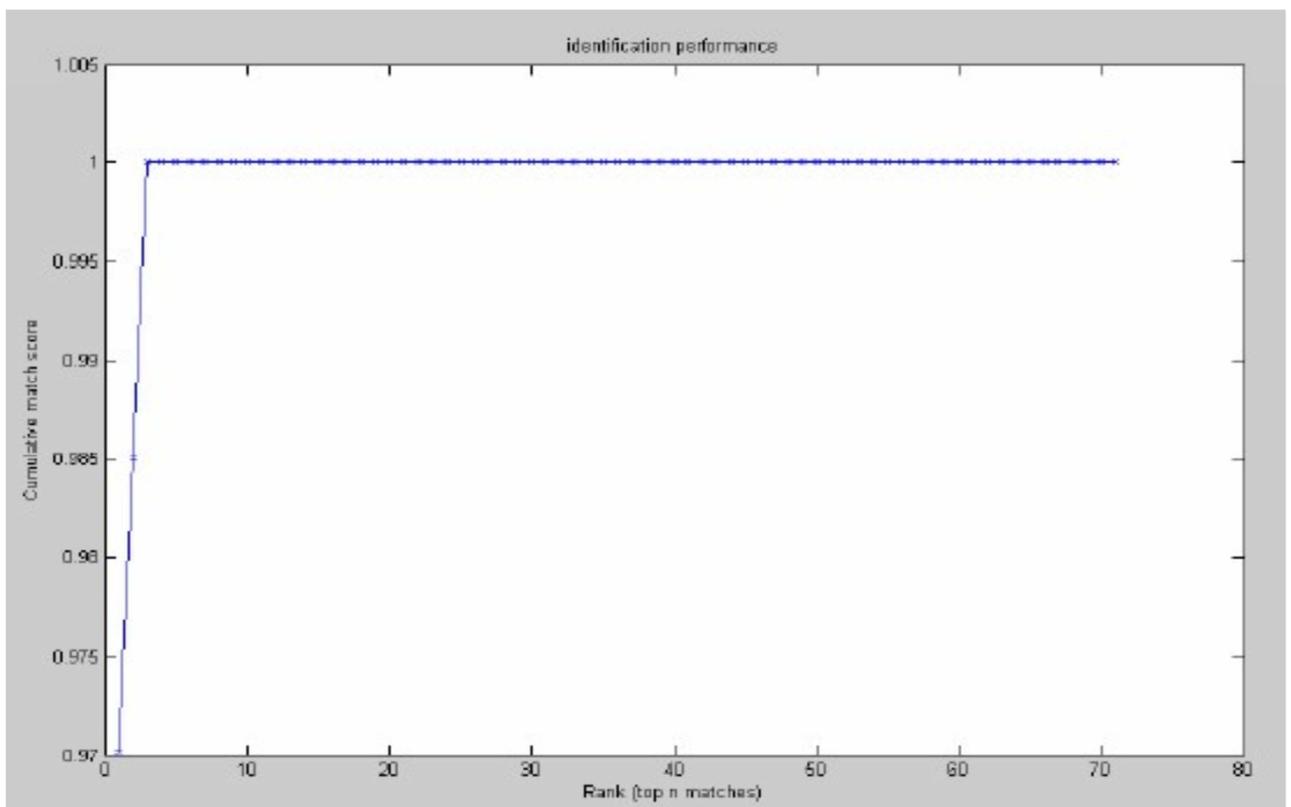
**FIG. B.10-** Le graphe des performances d'identification de PCA sur le Set4 de la base HARVARD



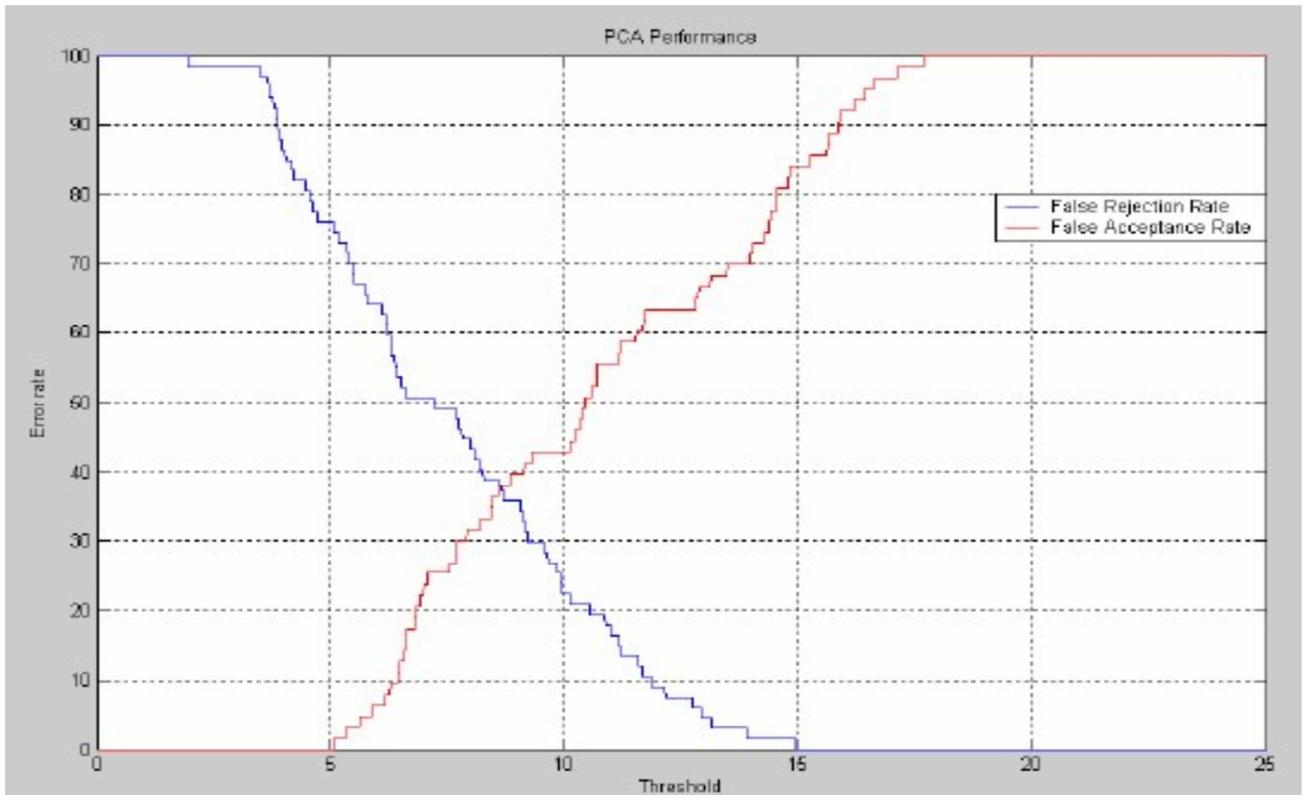
**FIG. B.11-** Le graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil, de PCA sur le Set4 de la base HARVARD



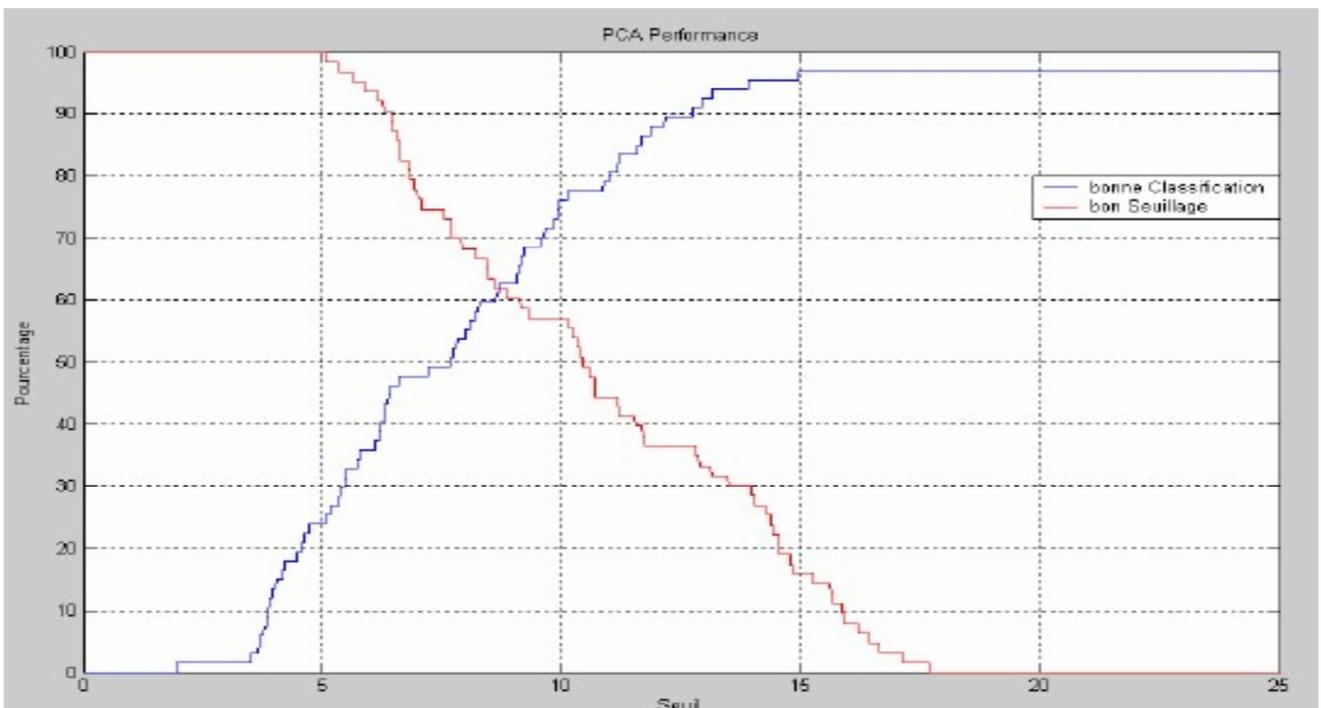
**FIG. B.12-** Le graphe des taux du bon seuillage et de la bonne classification en fonction de la valeur du seuil, de PCA sur le Set4 de la base HARVARD



**FIG. B.13-** Le graphe des performances d'identification de PCA sur le Set5 de la base HARVARD



**FIG. B.14-** Le graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil, de PCA sur le Set5 de la base HARVARD



**FIG. B.15-** Le graphe des taux du bon seuillage et de la bonne classification en fonction de la valeur du seuil, de PCA sur le Set5 de la base HARVARD

Annexe C Les graphes des performances d'identification et du seuillage de PCA, avec 10 et 15 d'images d'apprentissage par personne

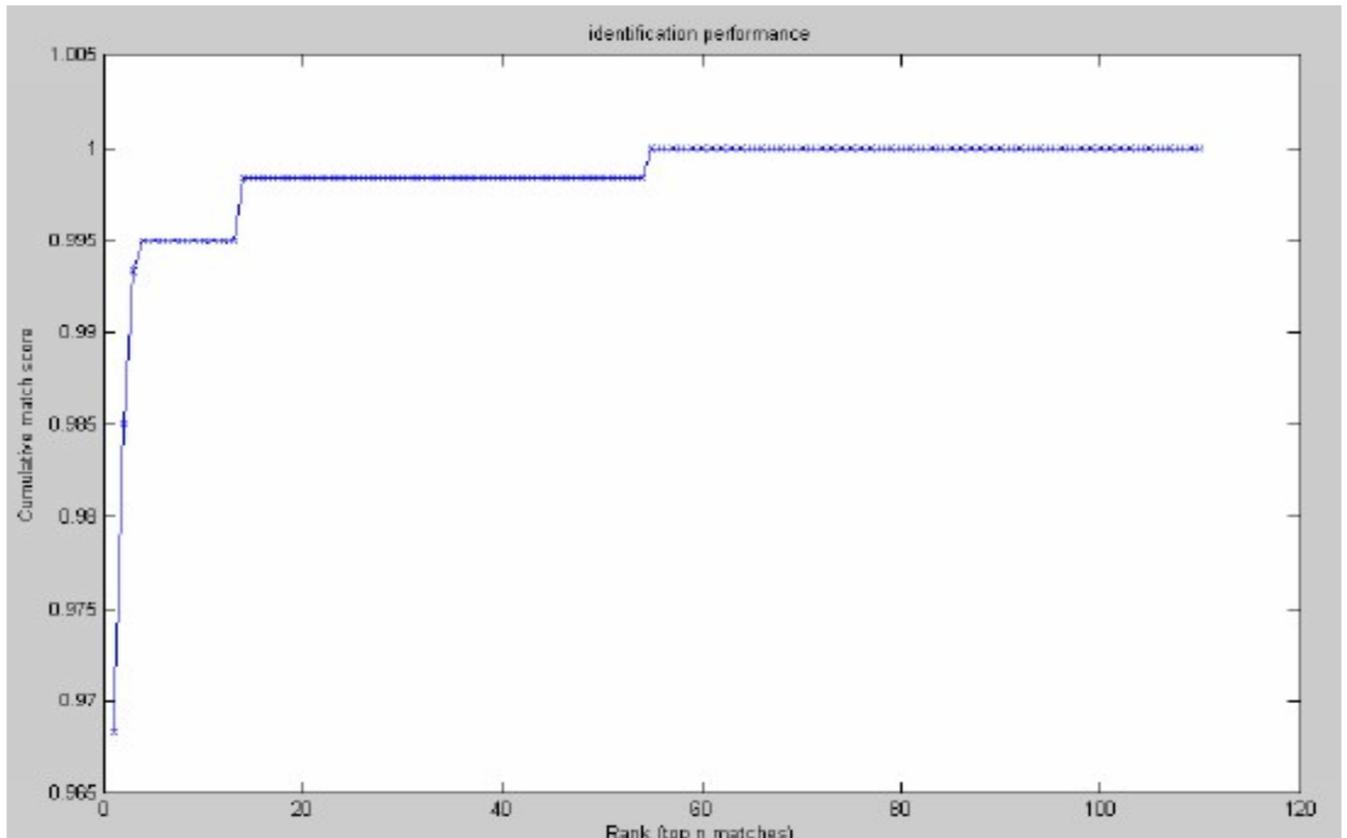


FIG. C.1- Le graphe des performances d'identification de PCA sur UMIST, en utilisant 10 images d'apprentissage par personne

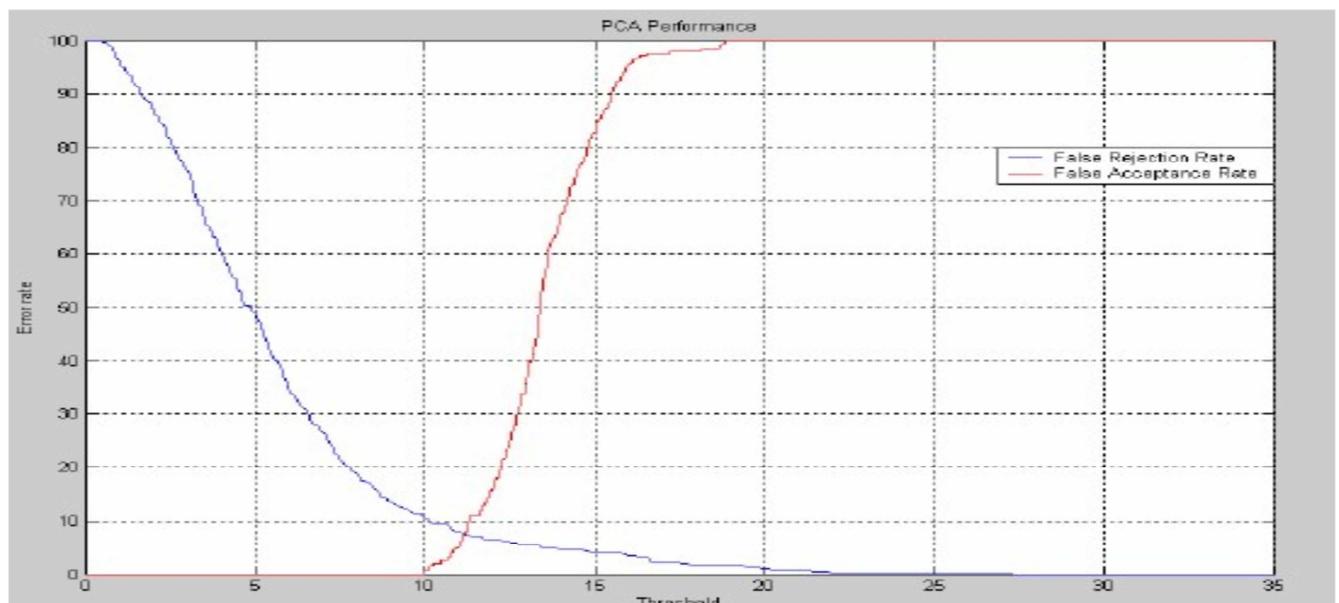
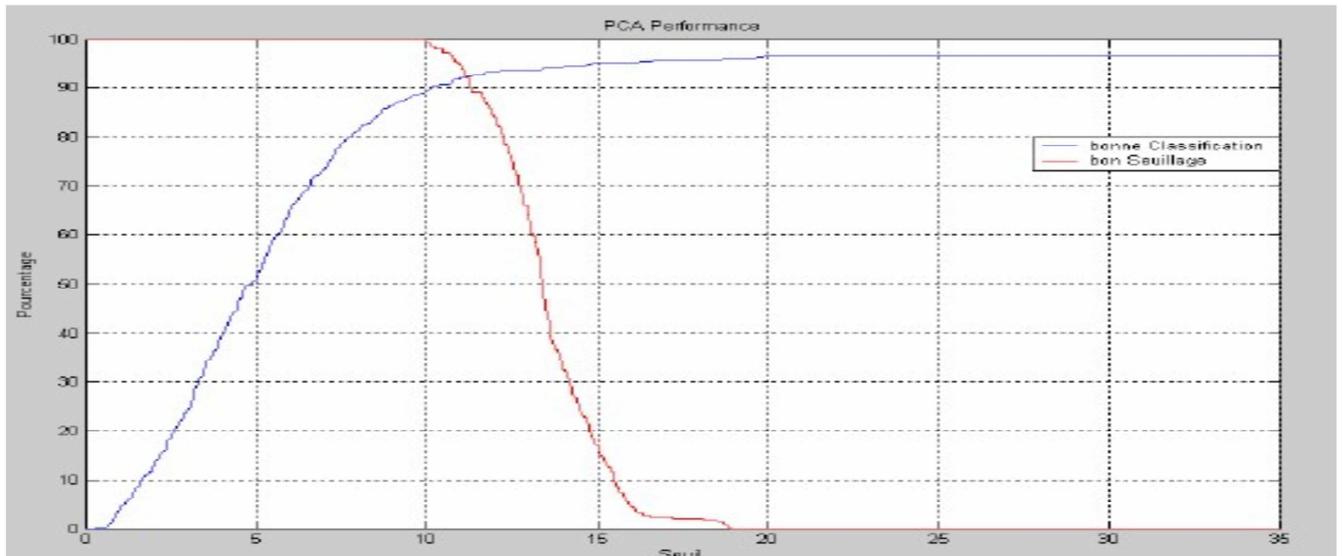
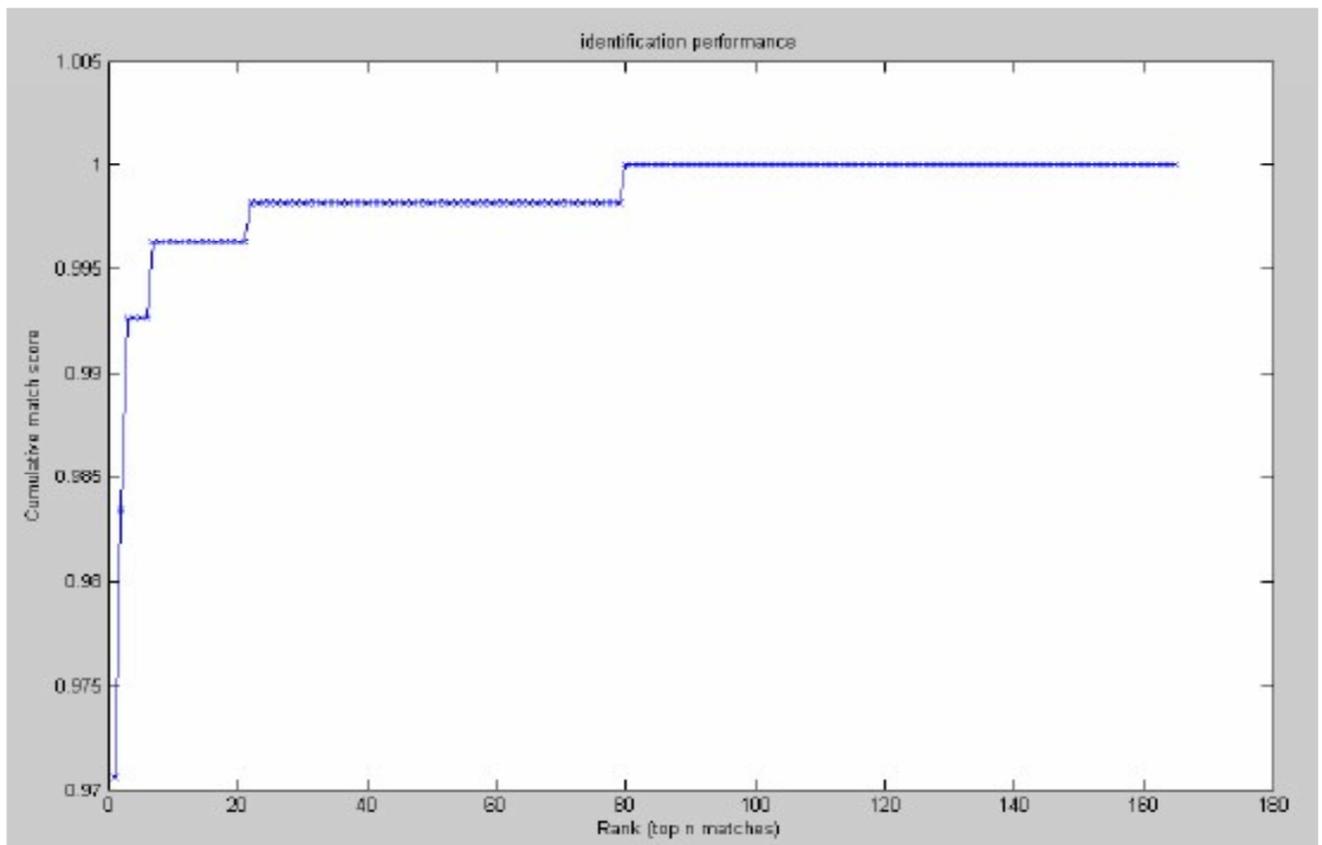


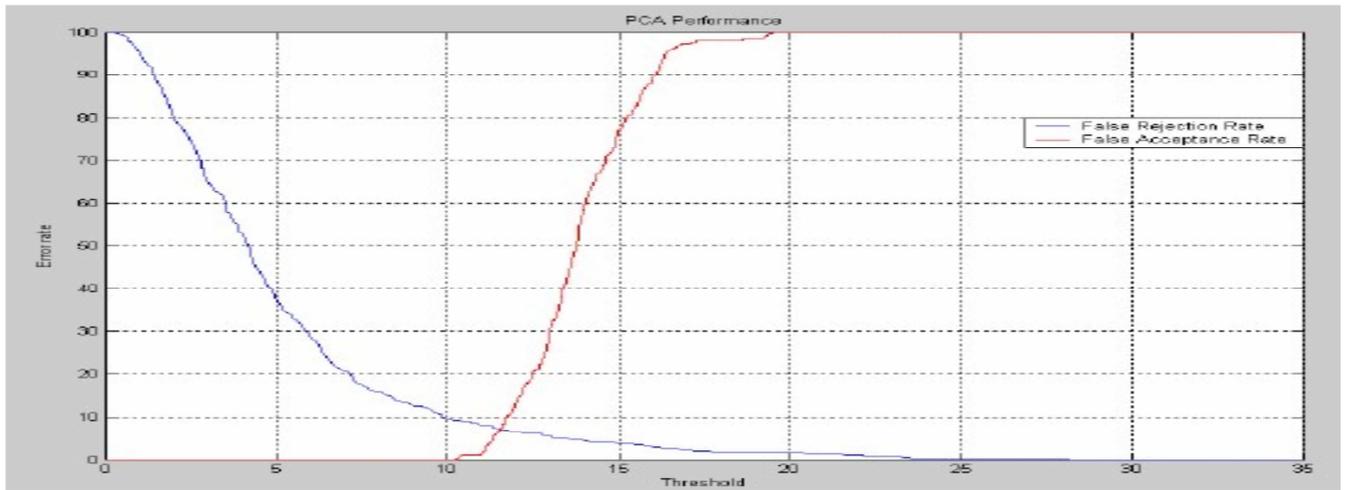
FIG. C.2- Le graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil, de PCA sur UMIST, en utilisant 10 images d'apprentissage par personne



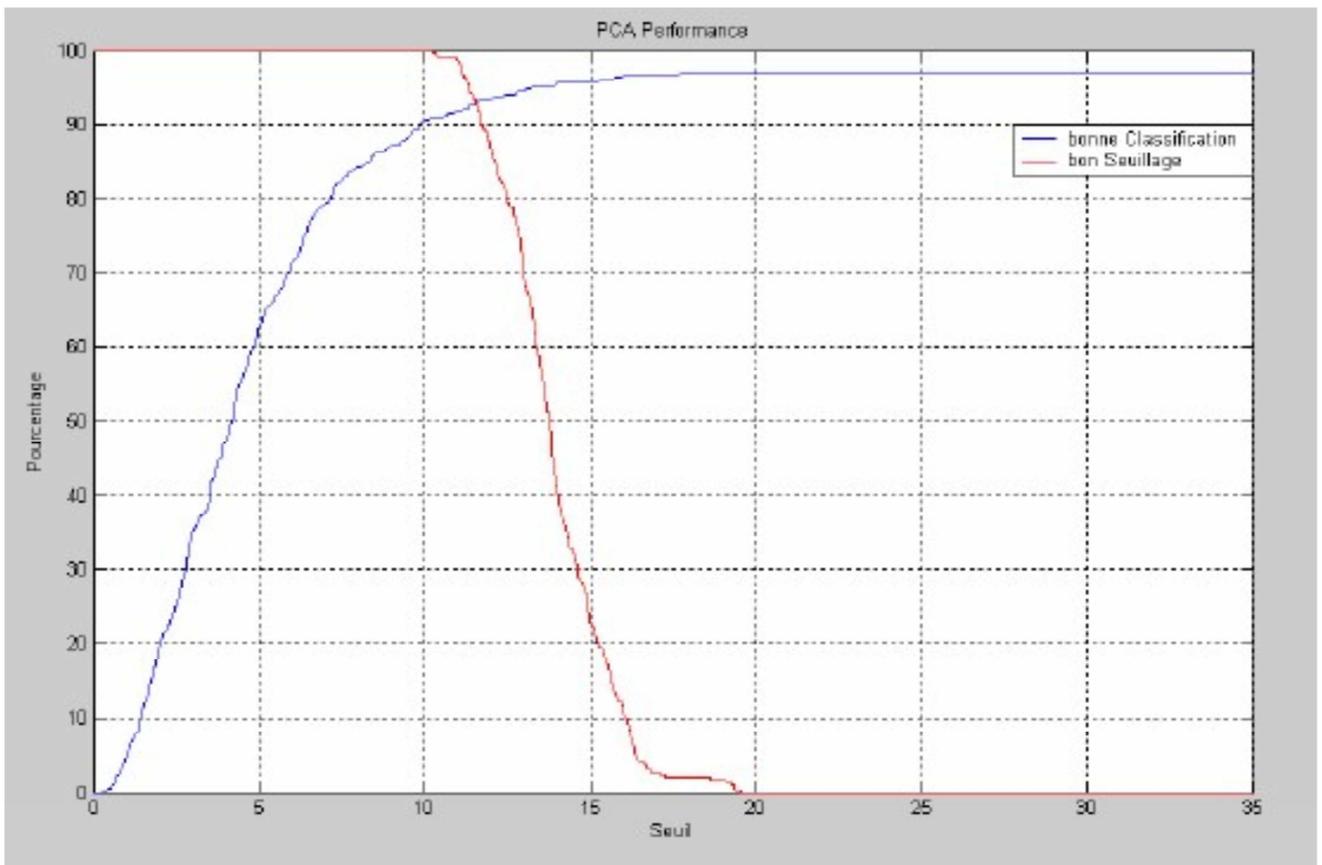
**FIG. C.3-** Le graphe des taux du bon seuillage et de la bonne classification en fonction de la valeur du seuil, de PCA sur UMIST, en utilisant 10 images d'apprentissage par personne



**FIG. C.4-** Le graphe des performances d'identification de PCA sur UMIST, en utilisant 15 images d'apprentissage par personne



**FIG. C.5-** Le graphe des valeurs de FRR et de FAR en fonction de la valeur du seuil, de PCA sur UMIST, en utilisant 15 images d'apprentissage par personne



**FIG. C.6-** Le graphe des taux du bon seuillage et de la bonne classification en fonction de la valeur du seuil, de PCA sur UMIST, en utilisant 15 images d'apprentissage par personne

## Prérequis mathématiques

### 1.1 Statistiques

#### 1.1.1 Moyenne

On considère un vecteur  $X$  représentant une distribution de  $n$  données :  $X = [X_1, X_2, \dots, X_i, \dots, X_n]$ .

La moyenne  $X'$  de cette distribution s'écrit :

$$X' = \sum_{i=1}^n \frac{X_i}{n} \quad (1)$$

#### 1.1.2 Écart-type

L'écart-type est une mesure de la dispersion d'un ensemble de données. D'un point de vue qualitatif, l'écart-type caractérise la largeur d'une distribution de données en mesurant la dispersion autour de la moyenne.

La formule de l'écart-type  $\sigma$  est :

$$\sigma = \sqrt{\sum_{i=1}^n \frac{(X_i - X')^2}{n-1}} \quad (2)$$

*Remarque : On se pose généralement la question ((Pourquoi  $n-1$  ?)). La raison pour laquelle on divise par  $n - 1$  au lieu de  $n$  est issue de l'interaction permanente entre les statistiques et les probabilités.*

#### 1.1.3 Variance

La variance est une autre mesure de la dispersion d'un ensemble de données. Il s'agit tout simplement du carré de l'écart-type.

La formule de la variance  $\sigma^2$  est donc :

$$\sigma^2 = \sum_{i=1}^n \frac{(X_i - X')^2}{n-1} \quad (3)$$

### 1.1.4 Covariance

La covariance est une extension de la notion de variance. Elle mesure la variation simultanée de deux variables et donne une matrice carrée. La formule de covariance entre deux variables X et Y s'écrit donc :

$$Cov(X, Y) = \sum_{i=1}^n \frac{(X_i - \bar{X})(Y_i - \bar{Y})}{n-1} \quad (4)$$

Si les quantités  $(X_i - \bar{X})$  et  $(Y_i - \bar{Y})$  varient dans le même sens (on dit alors que les variables X et Y covariant), la covariance sera positive. À l'inverse, si  $(X_i - \bar{X})$  et  $(Y_i - \bar{Y})$  varient dans des sens opposés, la covariance sera négative. Enfin, on voit bien d'après (4) que si la covariance est calculée entre une variable et elle-même (donc si  $X = Y$ ), nous obtenons (3), donc la variance.

### 1.1.5 Matrice de Covariance

La matrice de covariance permet de mesurer la variation simultanée de n variables. Elle nous sera très utile puisque nous serons amenés à manipuler des vecteurs avec de très nombreuses composantes, lors des algorithmes 2D de reconnaissance du visage. Pourquoi utiliser une matrice ? Si nous devons travailler dans un espace vectoriel de dimension supérieure à 2, plusieurs mesures de covariance peuvent alors être calculées. Plus précisément, pour n dimensions, nous pouvons calculer  $\frac{n!}{2*(n-2)!}$  valeurs différentes de covariance (en excluant les termes de la diagonale principale). Une manière pratique d'obtenir toutes ces valeurs est de les calculer et de les mettre dans une matrice. Ainsi, pour n dimensions, la définition d'une matrice de covariance est :

$$C^{n \times n} = (C_{i,j}, C_{i,j} = cov(Dim i, Dim j))$$

Où  $\begin{cases} C^{n \times n} \text{ Matrice avec } n \text{ ligne et } n \text{ colonne} \\ Dim x \text{ la } x^e \text{ dimension} \end{cases}$

Par exemple, en dimension 3, utilisant les dimensions usuelles x,y,z, nous obtenons une matrice de covariance de taille  $(3 \times 3)$  et dont les valeurs sont les suivantes :

$$C = \begin{pmatrix} cov(x,x) & cov(x,y) & cov(x,z) \\ cov(y,x) & cov(y,y) & cov(y,z) \\ cov(z,x) & cov(z,y) & cov(z,z) \end{pmatrix}$$

Remarque : le long de la diagonale, on voit que la valeur de la covariance est selon une dimension et elle-même. Ce sont les variances pour ces dimensions. L'autre chose à noter est que comme  $cov(a, b) = cov(b, a)$ , la matrice est symétrique par rapport à sa diagonale principale.

## 1.2 Algèbre linéaire

Dans cette section, nous allons revoir les notions de vecteurs propres et de valeurs propres qui jouent un rôle essentiel dans les algorithmes globaux.

### 1.2.1 Vecteurs propres et valeurs propres

Comme nous le savons, nous pouvons multiplier deux matrices entre-elles, sous réserve de compatibilité de leurs dimensions respectives. Les vecteurs propres sont un cas particulier de cette opération matricielle.

Soit E un vecteur propre d'une matrice carrée M, et  $\lambda$  un réel, alors :

$$M \cdot E = \lambda \times E, \lambda \in \mathbb{R} \quad (6)$$

Où  $\lambda$  est la valeur propre associée au vecteur propre E. Plus la valeur propre est grande, plus l'information contenue dans le vecteur propre correspondant est importante.

Exemple :

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} * \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 12 \\ 8 \end{pmatrix} = 4 * \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

Où :

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \text{ La matrice M}$$

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} \text{ Le vecteur propre E}$$

4 la valeur propre  $\lambda$

# Bibliographies:

- [1] John D. Woodward, Jr., Christopher Horn, Julius Gatune, and Aryn Thomas,  
“Biometrics A Look at Facial Recognition”
  
- [2] “Biometric system – IDTECK”.  
<http://www.idteck.com>
  
- [3] Mohamed Tayeb Laskri and Djallel Chefrour, “Who\_Is: système d’identification  
des visages humains”, A R I M A – Volume 1 – 2002
  
- [4] Dr. Andrzej Drygajlo, ELE 233, “BIOMETRICS”.  
<http://scgwww.epfl.ch>
  
- [5] Wen Gao, Shiguang Shan, Face Verification for Access Control, “Biometrics  
Solutions for Authentication in an E-World”
  
- [6] W. Zhao, R. Chellappa, P. J. Phillips and A. Rosenfeld, “Face Recognition: A  
Literature Survey”
  
- [7] Xiaoyin xu, “Image Based Face Recognition using Global Features”  
<http://www.vlsi.uwindsor.ca/>

- [8] Damien Dessimoz, Jonas Richiardi, Prof. Christophe Champod and Dr. Andrzej Drygajlo, “Multimodal Biometrics for Identity Documents State-of-the-Art”, Report de recherche, PFS 341-08.05, (Version 1.0). September 2005.
- [9] Lisa Osadciw, Pramod Varshney, Kalyan Veeramachaneni, “OPTIMUM FUSION RULES FOR MULTIMODAL BIOMETRIC SYSTEMS
- [10] “Principal components analysis”.  
[http://en.wikipedia.org/wiki/Principal\\_component\\_analysis](http://en.wikipedia.org/wiki/Principal_component_analysis)
- [11] A. S. Tolba, A.H. El-Baz, and A.A. El-Harby, “Face Recognition: A Literature Review”
- [12] Cherng Jye Liou, “A Real Time Face Recognition System”
- [13] Alex Chirokov, “Karhunen-Loeve, for face recognition  
<http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=6995&objectType=file>.
- [14] “Linear discriminant analysis”.  
[http://en.wikipedia.org/wiki/Linear\\_discriminant\\_analysis](http://en.wikipedia.org/wiki/Linear_discriminant_analysis).

- [15] Juwei Lu, Kostantinos N. Plataniotis, and Anastasios N. Venetsanopoulos, “Face Recognition Using LDA-Based Algorithms
- [16] M. Bartlett and H. Lades and T. Sejnowski, “Independent component Representations for face recognition”
- [17] “Modules complémentaires à MATLAB”.  
<http://www.ryounes.net>
- [18] L. Png, “Morphological Shared-Weight Neural Network For Face Recognition”
- [19] “Support vector machine”.  
[http://en.wikipedia.org/wiki/Support\\_vector\\_machine](http://en.wikipedia.org/wiki/Support_vector_machine).
- [20] “Cours IFT6266, SVM: support vector machines”.  
<http://www.iro.umontreal.ca/~bengioy/ift6266/svm/svm.html>.
- [21] Bernd Heisele, Purdy Ho, Jane Wu, and Tomaso Poggio, “Face recognition: component-based versus global approaches”
- [22] “Visual Object Recognition using Template Matching”.  
<http://www.araa.asn.au/acra/acra2004/papers/cole.pdf>.
- [23] Gareth James, “Investigation into Matching High Level Features Across Images”,

thèse présentée pour l'obtention de BSc (Hons) en sciences informatiques de l'Université Bain de, Juin 2005.

[24] “Distance de Hausdorff”.

[http://fr.wikipedia.org/wiki/Distance\\_de\\_Hausdorff](http://fr.wikipedia.org/wiki/Distance_de_Hausdorff).

[25] Daniel B Graham and Nigel M Allinson, “Characterizing Virtual Eigensignatures For General Purpose Face Recognition”, Face Recognition (en): De la théorie aux applications, NATO ASI Série F, de l'informatique et des sciences des systèmes, vol. 163.H. Wechsler, PJ Phillips, V. Bruce, F. Fogelman-Soulie et TS Huang (eds)

[26] “Resources for Face Detection”.

<http://vision.ai.uiuc.edu/mhyang/face-detection-survey.html>.

[27] The Color FERET Database's main webpage:

<http://www.nist.gov/humanid/colorferet>

[28] The Color FERET DVD's .

[29] P. Jonathon Phillips and Hyeonjoon Moon and Syed A. Rizvi and Patrick J. Rauss, “The FERET Evaluation Methodology for Face-Recognition Algorithms”, IEEE Transactions on Pattern Analysis and Machine Intelligence, volume 22, number

10

- [30] Matlab code and data  
<http://bach.ece.jhu.edu/pub/gert/svm/incremental>.
- [31] G. Cauwenberghs and T. Poggio, “Incremental and Decremental Support Vector Machine Learning” in Adv. Neural Information Processing Systems (NIPS\*2000), Cambridge MA: MIT Press, vol. 13, 2001.
- [32] Alan Brooks (in collaboration with Li Gao), “Face Recognition: Eigenface and Fisherface Performance Across Pose”, Vision de la CEE ordinateur 432 avec le Professeur Ying Wu, Rapport final du projet soumis le 9 Juin 2004.
- [33] Duda et al., “Bayesian Decision Theory”, Chapter 2, CS479/679 Pattern Recognition Spring 2006
- [34] The FastICA MATLAB package.  
<http://www.cis.hut.fi/projects/ica/fastica/>.
- [35] W. Zhao and R. Chellappa, “Image-based Face Recognition: Issues and Methods”, In Pattern Recognition: Algorithm, Systems and Applications, (B. Javidi, ed.), Marcel Dekker, Inc., New York, NJ, 2002.
- [36] “FaceVACS-SDK Version 4.0.3”.  
[http://www.cognitec-systems.de/documentation/frsdk\\_public/index.html](http://www.cognitec-systems.de/documentation/frsdk_public/index.html)

[37] <http://www.irit.fr/page-perso/RJ%20-%20DESA%20IT.pdf>